REC'D NOV 20 1961

December, 1961
Vol. 32, No. 4

Child Development

CONTENTS

Family Role Concepts of Children Ages Six to Ten
Walter Emmerich 609

Child's Symbolic Conceptualization of Parents
Jerome Kagan, Barbara Hosken, and **Sara Watson** 625

Assessing Some Aspects of the Parent-Child Relationship
F. N. Cox and **Patricia M. Leaper** 637

Adaptability: A Problem Solving Approach to Assessing Childrearing Practices
Irving Tallman 651

Childrearing Attitudes of Mothers in Germany and the United States
Don W. Rapp 669

New Techniques for Research in Children's Temptation Behavior
Robert E. Grinder 679

Antecedents of Resistance to Temptation in Four-Year-Old Children
Roger V. Burton, Eleanor E. Maccoby, and **Wesley Allinsmith** 689

Milieu Influences in Children's Concepts of Misconduct
Paul V. Gump and **Jacob S. Kounin** 711

Measurement of Children's Sensitivity to Audiences
Allan Paivio, Alfred L. Baldwin, and **Seymour M. Berger** 721

Personality and Behavior Patterns of Children with Functional Defects of Articulation
Arthur L. Solomon 731

Effects of Pretraining on the Reinforcing Value of Visual Stimuli
Harold W. Stevenson and **Richard D. Odom** 739

Effects of Pretraining with Varied Stimuli on Children's Discrimination Learning
Sheldon H. White 745

Secondary Reinforcement with Preschool Children
Jane G. Fort 755

Secondary Reinforcement as a Function of the Number of Reinforced Trials
Nancy A. Myers, Grace J. Craig, and **Jerome L. Myers** 765

Effects of Associative Strength, Grade Level, and Interpair Interval in Verbal Paired-Associate Learning
John C. McCullers 773

Behavior of Young Children in a Two-Choice Guessing Problem
William Kessen and **Marion Lord Kessen** 779

Effects of Delayed Reinforcement on Associative and Incentive Factors
Robert Ware and **Glenn Terrell** 789

Published Quarterly by Child Development Publications of the

CIETY FOR RESEARCH IN CHILD DEVELOPMENT, INC.

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

DONALD M. BAER
University of Washington

EMANUEL K. BELLER
Child Development Center

ALFRED CASTANEDA
University of Texas

WALTER EMMERICH
University of Colorado

JEROME KAGAN
Fels Research Institute

IRVIN A. KRAFT
Houston State Psychiatric Institute

LEWIS P. LIPSITT
Brown University

NORMAN LIVSON
University of California

JAMES M. TANNER
University of London

RICHARD H. WALTERS
University of Toronto

Managing Editor

KATE HOFFMAN

Child Development, one of three publications of the Society for Research in Child Development, Inc., is issued quarterly in March, June, September, and December. The subscription price per year is $12.00 domestic, $12.50 foreign. Single issues are $3.50.

Child Development Abstracts and Bibliography is issued three times a year, two numbers in each issue. The subscription price per year is $6.00 domestic, $6.50 foreign. Single issues are $2.50.

Monographs of the Society for Research in Child Development is issued at irregular intervals during the year. The subscription price per year is $6.00 domestic, $6.50 foreign. The price of individual monographs is variable.

The three publications of the Society are available at the special price of $20.00 per year, domestic; $21.50 per year, foreign.

A limited number of back issues of all publications is available. A mimeographed list may be had upon request.

Subscriptions, address changes, and business communications should be sent to CHILD DEVELOPMENT PUBLICATIONS, Purdue University, Lafayette, Indiana.

Please give four weeks' notice when changing your address, giving old and new addresses. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Printed and mailed by
THE ANTIOCH PRESS
Yellow Springs, Ohio

FAMILY ROLE CONCEPTS OF CHILDREN AGES SIX TO TEN[1,2]

WALTER EMMERICH*
University of Colorado

A number of social scientists have emphasized the importance of role perception as a determinant of social action (1, 9, 15, 18, 20). It is believed that role discriminations help define the constancies anticipated in each person's behavior during the ensuing interaction. Insofar as such anticipations function to mediate and guide purposeful activity, a thorough understanding of their nature is essential for predicting behavior. Very little is known about the dimensions children use to define dyadic relationships within the family. In an earlier study by the writer (7), this problem was investigated in children of 4½ to 6 years. The present study extends the investigation to an older sample, with the purpose of describing intragroup role consensus and differences in consensus among subgroups classified according to sex and age.

The descriptive aim of this study is pertinent to several areas of developmental social psychology. First, any subgroup regularities in role perception will in part describe the typical child's social space and thereby become parameters in predicting his behavior. Second, the functioning of a social system is presumably dependent upon the extent of agreement on role definitions among its members. By comparing normative data on children and parents, it should be possible to determine characteristic areas of definitional agreement and disagreement between parent and child and to relate these to indices of family integration. Third, sex and age variations in role consensus have bearing on theories of developmental change in the cognition

* Department of Child Development and Family Life, Purdue University, Lafayette, Indiana.

1 This investigation was supported by research grants M-2853 (A) and M-3844 (A) from the National Institute of Mental Health, United States Public Health Service.

2 The writer wishes to express his appreciation to Marilyn McRae and Ronald Cunat for their assistance in this study and to Assistant Superintendent Melvin Wiesley, Principal Keith Dillman, and the teachers of the Martin Park Elementary School, Fairview District, Boulder, Colorado, whose cooperation made this study possible.

Child Develpm., 1961, 32, 609-624.

of social structures, such as that of Parsons (18). Fourth, if role conceptions are influenced by cultural arrangements which vary with such factors as social status, sex, and age, it becomes necessary to have subgroup norms in order to evaluate properly the degree of "normalcy" of a particular individual's role perceptions. Finally, as Parsons (18) and Brim (4) point out, knowledge about specific intrafamilial role relationships is essential for understanding the precise nature of more complex role processes. To illustrate this point, consider the commonly held belief that sex-identity is ordinarily acquired by the child's taking one of his parents as a model. Assuming that this belief is correct,[3] the problem still remains of specifying which role relationships serve as referents of the modeling process. For example, in the case of the girl who takes her mother as a model in play, to what extent does her play behavior reflect the mother's relationship with the girl, the father, a sibling, a person outside the family, or some combination of these? Suppose that the girl discriminates age roles, but not sex roles. Then she might take the mother-girl relationship as a model, but she cannot do so for the mother-father relationship because she does not discriminate sex roles. The assessment of children's role concepts, then, is believed to be prerequisite to the analysis of identification patterns acquired through a process of role modeling.[4]

It is important to select for study those role dimensions which are likely to be salient for both sexes over an extended age period. The literature indicates that *relative power* and *attitudinal direction* are two dimensions which meet this criterion (3, 8, 13, 14). Relative power is defined here as the amount of implied or explicit control each person has over the outcome of an interaction. Attitudinal direction is defined as the extent of implied or explicit agreement (positive) or disagreement (negative) with respect to the means and/or goals of an interaction. Other definitions used in the study are as follows. A *role* is defined as the behaviors (performances) characteristic of the incumbent of one position toward the incumbent of another position in a social structure. This definition is similar to Gross, Mason, and McEachern's (9) concept of "role behavior sector." A *role concept* is defined as the discrimination of a particular position on a specific behavioral dimension. The referent of a *person concept* is distinguished from that of a *role concept.* A person concept is derived from a particular person's performances of a variety of roles, whereas a role concept is derived from the performance of a typical person or many persons in a single role. There are at least two criteria for operationally distinguishing person and role concepts. For a role, but not for a person concept, there is: (a) generalization to the population of persons who hold the same position; (b) intragroup

[3] Bronfenbrenner (5) points out that there is little direct support for this assumption.

[4] This discussion assumes that the child's role concepts and models are assessed from the child's reports, rather than from observations of overt parent and child behaviors. Nevertheless, the point made here would also apply to the latter case, since family roles must in fact differ if differential modeling is to occur.

consensus on the behavior characteristic of a position, when each group member judges a different occupant of the position. This study relies primarily upon the second of these criteria.

Method

Subjects

Subjects were 225 elementary school children from middle-class families unbroken by separation, death, or divorce. None was adopted. Age groups were formed in such a way as to create nonoverlapping age ranges, approxi-

Table 1

DESCRIPTION OF SUBJECTS

		Age (in Months)		
	N	*Mean*	*Range*	*SD*
6 Year Group				
Girls	21	71.5	66– 75	2.7
Boys	17	70.4	64– 75	3.4
Total	38	71.0	64– 75	3.1
6½ Year Group				
Girls	18	78.9	76– 82	2.0
Boys	19	79.7	76– 82	1.8
Total	37	79.3	76– 82	1.9
7 Year Group				
Girls	19	87.9	83– 91	2.8
Boys	20	86.8	83– 91	2.9
Total	39	87.3	83– 91	2.8
8 Year Group				
Girls	17	96.2	92–104	3.7
Boys	21	96.4	92–103	3.2
Total	38	96.3	92–104	3.7
9 Year Group				
Girls	20	108.7	105–116	3.7
Boys	16	108.7	104–115	3.5
Total	36	108.7	104–116	3.5
10 Year Group				
Girls	16	121.1	117–131	3.7
Boys	21	121.6	117–129	3.3
Total	37	121.4	117–131	3.4
Total				
Girls	111	93.2	66–131	17.2
Boys	114	94.4	64–129	17.6
Total	225	93.8	64–131	17.4

mately the same mean age for both sexes within each age level, and approximately equal *N*s in all subgroups. Six age groups resulted, henceforth called the 6, 6½, 7, 8, 9, and 10 year groups. Subgroup sizes and age statistics are given in Table 1. Approximately 79 per cent of the sample came from socioeconomic levels 1 or 2 in Warner *et al.*'s (occupational) classification scheme (25), and approximately 76 per cent of the sample had either one or two siblings.[5]

Materials

Stylized figure drawings, described in detail elsewhere (7), were prepared on 6¼ by 3¼ in. Hunter Card Master cards in the following pairs: mother-girl, father-boy, mother-father, and girl-boy. In order to mix the presentation of pairs and items, 16 cards were made for the mother-girl and father-boy pairs, and eight cards were made for the mother-father and girl-boy pairs. Each member of a pair appeared on the left and right sides of the cards half the time. The 48 cards were presented to subjects by means of the Hunter Card Master, Model 340.

Procedure

Subjects were tested individually. *S* sat in front of the exposure window of the Card Master, with *E* seated beside him. *E* pointed to the exposed boy-father card and said, "Do you see these people here? Let's play a game with them. Okay? Who says, 'I'm the boy'?" After *S* pointed to the correct figure, *E* said, "Yes, that's a boy just like you" if *S* was a boy, or "Yes, that's a boy just your age," if *S* was a girl. *E* then asked, "Who says, 'I'm the father'?" After *S* pointed to the correct figure, *E* exposed the next card which was the girl-mother pair. Experimenter then asked, "Who says, 'I'm the girl'?" After *S* pointed to the correct figure, *E* said, "Yes, that's a girl just like you," if *S* was a girl, or, "Yes, that's a girl just your age," if *S* was a boy. *E* then asked, "Who says, 'I'm the mother'?" In order to give some initial practice with behavior descriptions similar to the actual test items, the girl-boy card was presented next, with the item, "I want it." The mother-girl card was then presented with the item, "Stop doing that." The test proper then began and consisted of presenting the 48 pairs in a prearranged order, designed to control for a variety of possible response sets. Either pointing to a figure or naming it was an acceptable response. *S* was given as much time as he wanted, and *E* repeated the item whenever *S* did not respond after several seconds or responded ambiguously. Each child was told that he did very well at the completion of testing.

In order to have partial control over the influences of examiner characteristics on the children's responses, a male and a female tested approximately half the subjects in each of the 12 subgroups.

---

[5] These figures are based upon the 157 cases for which exact data were available on socioeconomic level and number of siblings. The remaining 68 subjects were very similar to this group in both respects.

TABLE 2

CLASSIFICATION OF ITEMS ACCORDING TO RELATIVE POWER AND ATTITUDINAL DIRECTION

High Power	*Low Power*
Positive Attitude	
1. You can have it.	1. I want it.
2. That's nice. You did what I asked.	2. I'll do what you say.
3. You made that very well.	3. I made it very well.
4. You did what was right.	4. I'll do what is right.
Negative Attitude	
1. You can't have it.	1. Give me what *I* want.
2. You'd better do as I say.	2. No. I won't do it.
3. You didn't make it very well.	3. I can't make it very well.
4. You did something that is wrong.	4. I'll do something that is wrong.

Measures

The items were similar to those used in the previous study (7) and are classified in Table 2 according to relative power and attitudinal direction. Items having corresponding numbers across columns and rows are role reciprocals on the power and attitude dimensions, respectively. All items were used to assess age-role discriminations of the mother-girl and father-boy pairs. Because of limited time for testing and because of the tendency for younger subjects to assign high and low power items to the adult and child generations, respectively (7), high power items only were used to assess sex-role discriminations of the mother-father pair, while low power items only were used to assess sex-role discriminations of the girl-boy pair. For the mother-girl and father-boy comparisons, the ***power discrimination measure*** was the total number of assignments of high power items to adult figures and low power items to child figures. For the mother-father comparison, this measure was the number of times the father was chosen on high power items, and, for the girl-boy comparison, this measure was the number of times the girl was chosen on low power items. For the mother-girl and father-boy comparisons, the ***attitude discrimination measure*** was the total number of assignments of positive items to adult figures and negative items to child figures. For the mother-father and girl-boy comparisons, this measure was the total number of assignments of positive items to female figures and negative items to male figures.

The above measures were based upon the assumption that all items contributed to both dimensions according to the scheme presented in Table 2. A number of scale characteristics, such as small variances and skewed dis-

tributions, made it impossible to estimate scale homogeneities and the discrimination power of items with precision. However, crude analyses of part-whole association were performed, using the extremes of those subgroups having maximal variability on the scales. All items were found to contribute to both scales applied to each of the four paired comparisons. The chi square probabilities associated with the 32 phi coefficients ranged from .20 to .001, with a median of .01.

Results and Discussion

Nonparametric statistics were used in all analyses, and a finding is interpreted if it was significant beyond the .05 level. For analyses of intragroup consensus, chi square was used to determine if more than half the *Ss* fell either above or below the scale point representing equal assignment of both ends of a dimension to each person in a pair. This point on the scale shall henceforth be referred to as "chance." The median test or its extension to several independent variables in a single analysis (26) was applied to differences among subgroups. The sign test was used in the analyses of differential item sensitivities. The index of association was *C* (significance levels only are reported), with distributions divided at or near the median.

Discrimination of Age Roles by Relative Power

There was a marked tendency for the whole sample to use the power dimension to discriminate the mother-girl and father-boy pairs. Scores for these two pairs combined could vary from 0 to 32, with 16 representing chance. Ninety-nine per cent of the whole sample had scores above 16, and the median was 29.5. It is clear that the power dimension is used by children of this age to discriminate age roles, as is the case for younger children (7). This result is highly consistent with Barker and Wright's (3) survey of actual parent-child interactions, in which it was found that dominance and nurturance are the most frequent parent behaviors, whereas appeal, submission, and resistance are the most frequent child behaviors.

Girls discriminated age roles on the power dimension somewhat more than boys ($p < .05$), and this difference did not vary with age. Although this finding is consistent with those suggesting that girls are more sensitive to age grading than boys (6, 10, 21) or that boys are more successful than girls in defying parental commands and restrictions (19), it does not very strongly support either of these interpretations, since boys clearly discriminated age roles by power, and the difference between the sexes was only about one point on the scale. Further, this sex difference did not appear in younger children (7).

Age changes in age-role discrimination by power are presented in Figure 1. The tendency for scores to increase with age is significant for the father-boy ($p < .001$), but not for the mother-girl comparison. There was no interaction between sex and age of child for either comparison. These find-

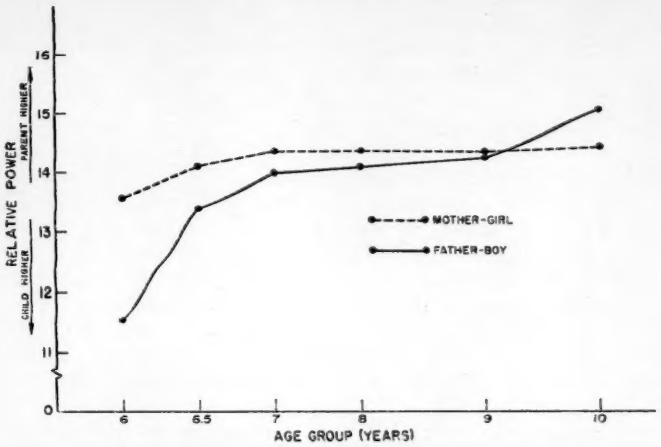

FIGURE 1—Discrimination of age roles by power.

ings are subject to several possible interpretations. If it is assumed that the age trend for the (nonassessed) father-girl comparison would have been similar to that found for the father-boy comparison, then the results signify that children age-type maternal relationships earlier than paternal relationships. Several factors could account for this possibility. Since mothers interact more frequently with children of this age than fathers (3), there is more opportunity for children to learn appropriate age roles in their maternal relationships. However, there is little evidence indicating that the frequency of father-child interactions increases with age during this period (3, 23), as did the father-boy discrimination. Or, if the mother is (or is perceived as) more nurturant or powerful than the father in socializing the younger child, then the child may be more motivated to overlearn appropriate age roles in the maternal relationship. Again, however, although some studies support this interpretation (3, 4, 19), others do not (4, 8, 13, 14). Another possibility is that the child's awareness of the father's high power in other roles increases with age and generalizes to the father-child relationship. This interpretation gains support from the finding, discussed in more detail later, that amount of power attributed to the father vis-à-vis the mother also increases rapidly between the sixth and eighth year (*see* Figure 2). On the other hand, if it is assumed that the age trend for the (nonassessed) mother-boy comparison would have been similar to that found for the father-boy comparison, then the findings signify that children of both sexes perceive that girls age-type earlier than boys. A number of studies do suggest that girls perceive, accept, or are expected to learn age-appropriate child and

adult roles earlier than boys (6, 10, 16, 19, 21). Further study is needed to determine which of these interpretations is most valid.

There was a greater tendency to assign high power items to parents than to assign low power items to children ($p < .001$), and this difference showed no consistent variation with sex or age. This is a replication of an earlier finding and is discussed elsewhere (7).

Parsons (18) has hypothesized that there are developmental changes in the specific cues used by the child to discriminate family roles. His discussion suggests the following developmental sequence of age-role discriminations: (a) *Response:* The child is primarily an object having needs; the parent is an agent of gratification or frustration of these needs. (b) *Acceptance:* The child is capable of meeting specific norms governing the parent-child relationship; the parent praises the child when he conforms, and controls the child when he deviates. (c) *Approval:* The child is capable of meeting universal standards of achievement; the parent responds positively to the child's successes and criticizes his failures. (d) *Esteem:* The child is capable of acting according to broad standards of conduct; the parent recognizes appropriate behavior and reproves inappropriate behavior.

The four kinds of items in each cell of Table 2 were formulated to correspond to each of these presumed developmental levels. It was hypothesized that age-role discrimination by power is a function of the interaction between item type and age of child; i.e., lower level items would be age-typed more by younger children, whereas higher level items would be age-typed more by older children. The results did not support this hypothesis, as age-role discrimination by power increased with age at about the same rate for all four types of items. Rather, certain items were more sensitive to this dimension, irrespective of the child's sex or age. Specifically, the scores on items 2 and 4 were significantly higher than those on items 1 and 3 ($p < .01$ for each of the four comparisons), and the differences between items within these pairs were not significant.

Do items 2 and 4 possess some quality lacking in items 1 and 3 which might account for the former pair's greater sensitivity to the power dimension? Whereas item 1 involves the exchange of an object and item 3 the exchange of meanings with reference to an objective standard of performance, items 2 and 4 seem to connote an exchange of affect directed toward the interacting persons themselves. If this analysis is correct, then it would suggest that children are more sensitive to personalized "love-oriented" than to "object-oriented" relationships (21) in their age-typing. Perhaps the kinds and/or patterning of rewards and punishments associated with love-oriented socialization techniques provide the child with greater incentive to learn age-role discriminations. This interpretation should be viewed with caution, however, for several reasons. Little is known about the differential incentive values of the rewards associated with each of the four types of items. Further, in view of the importance of achievement in middle-class society, particularly in boys, it is surprising that item 3 did not have stronger affective impli-

cations. Also, item affect saturation may vary among respondent populations, so that specification of affect criteria without reference to these populations would be misleading. It might be expected, for example, that item 1 would have stronger affective implications in a lower-class sample.

The failure of this study to confirm Parsons' hypothesis of developmental stages in children's age-role concepts should not be viewed as strong evidence against the theory. Parsons' discussion of levels (18) is not sufficiently detailed to be certain that the items used are highly representative of their respective level universes. Also, item differences on other dimensions, such as affect, may have masked the predicted interaction. Finally, since the hypothesis refers to changes in role perception which presumably begin during the first year, the present study may not have sampled age groups which are young enough for the effect to appear.

Discrimination of Age Roles by Attitudinal Direction

The median of the whole sample on the attitude age-role discrimination scale did not differ from chance. There were no significant variations from chance within any of the 12 subgroups, either for the separate mother-girl and father-boy comparisons or for these comparisons combined. It would appear that attitudinal direction is not a dimension in children's age-role discriminations. It should be noted, however, that the power and attitude discrimination scores for the same pair are not completely independent. The fact that the children used the power dimension so intensively and extensively made it less likely, though not impossible, for significant age-role discriminations to occur on the attitude dimension. This situation limits the interpretation, since an independent measure of age-role discrimination by attitude might have revealed trends masked by the power dimension. Nevertheless, if some subjects had relied heavily upon the attitude dimension, or if many had used this dimension a little more frequently, significant intragroup consensus or group differences could have appeared on this as well as on the power dimension. Therefore, it is reasonable to conclude that power is at least relatively more pertinent than attitude as a dimension of age-role discrimination during this age period.

Discrimination of Adult Sex Roles

The whole sample assigned more high power items to the father than to the mother ($p < .001$), and, as seen in Figure 2, there was a curvilinear relation between this tendency and age ($p < .05$). There was no interaction between sex and age on this measure. With respect to the attitude dimension, the median of the whole sample for mother-father discrimination did not differ from chance, and there was no significant variation with age. Although intragroup consensus on this measure was not significant for either sex, there was a difference between sex groups. Girls perceived the mother as somewhat more positive than the father, whereas boys perceived the converse ($p < .05$).

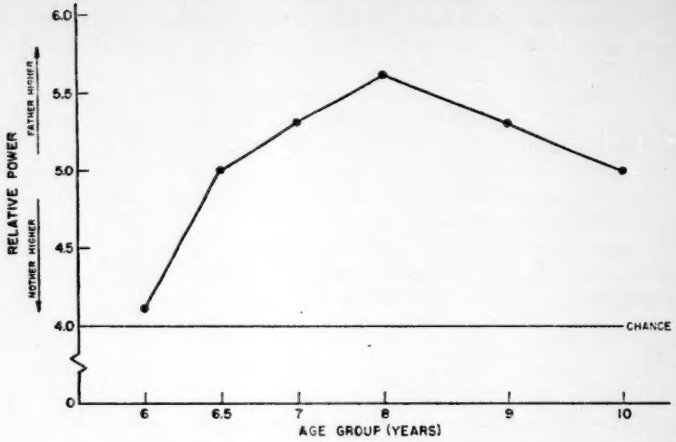

FIGURE 2—Discrimination of parent sex roles by power.

It is most striking that, for the sample as a whole, consensus again occurred on the power dimension, and not on the dimension of attitudinal direction. This finding, together with those on age-role discrimination, indicates that power is the more basic dimension in children's discriminations of parent roles, at least during this age period. If this conclusion is correct, then several interesting implications might follow. Although the importance of power has been recognized in the analysis of intrafamilial relationships (2, 3, 12, 14, 17, 18, 19, 22), many investigators, including the present writer, have assumed that the attitude dimension is at least equally important in parent role differentiation (4, 6, 7, 13, 14, 17, 18, 21). In view of the present results, the latter assumption seems questionable, at least for children of this age. It should be noted, however, that this conclusion does not imply lack of salience of the attitude dimension at other levels, such as that of *person perception,* but only that this dimension may not be pertinent at the level of *role consensus.* Indeed, other results of the present study suggest that several levels need to be considered in a theory of the perceptual factors which guide interpersonal behavior in the child. For example, consider the findings that both sexes perceive the father's sex role as more powerful than the mother's but that girls perceive the mother as more positive, whereas boys perceive the father as more positive. Heider's analysis of benefit and sentiment (11) suggests that a powerful person is seen as benevolent if he is also perceived as positive or as malevolent if he is also perceived as negative; whereas a nonpowerful person is differentiated less sharply in terms

of potential benefit. In the present example, it might follow that the father's sex role is perceived as more benevolent by boys and as more malevolent by girls, but that the mother's sex role is not perceived by either sex so clearly in terms of potential benefit or harm.[6]

A second implication is methodological and points out the value of deriving measures from a framework that explicates the various components of family interaction. By distinguishing two possible role dimensions (power and attitude) and two possible position dimensions (sex and age), it became possible to isolate critical role dimensions within specific intrafamilial dyads. Perhaps the often conflicting findings in this area, and consequent difficulties of interpretation, are partially due to inadvertent confounding of these and other interaction components.

The amount of power attributed to the father increased with age between the sixth and eighth year in both the father-boy and mother-father comparisons (Figures 1 and 2, respectively), indicating that this period is particularly significant in the development of father role concepts. Comparison of these two curves also suggests that while there is generalization across the sex- and age-role components of the father role during this period, differentiation of these role components begins to occur between the eighth and tenth year.

The pattern of results for the mother-father comparison differs in several respects from that found in the study of subjects ages 4½ to 6 (7). The younger sample as a whole discriminated parent sex roles on the attitude dimension, whereas the power dimension was used in the present older sample. The obvious interpretation of this difference is that there is a dimension crossover at about the age of 6. The fact that the 6-year-olds of the present study did not discriminate the mother-father pair on the power dimension is consistent with this interpretation. However, this group did not discriminate parent sex roles by attitude either. Further, the girls of the younger sample perceived the mother as more powerful than the father. It is difficult to account for these discrepancies on any basis other than age, since the procedures of the two studies were very similar and the subjects were selected from the same school in successive years. However, in view of the potential importance of a dimension crossover at about the age of 6, it would be desirable to replicate this aspect of the studies in a group ranging in age from 4 to 8 years.

It has been assumed throughout this discussion that the mother-father discrimination measure was derived independently of the measures of age-role discrimination discussed earlier. However, if upon presentation of the mother-father pair, the child imaginatively placed himself into the situation, then this measure could reflect differential parental behaviors toward the child, rather than toward each other. That such confounding may have

[6] This discussion is meant to provide an illustration, not specific hypotheses. Obviously, the findings on age-role discrimination would have to be incorporated into any set of actual predictions.

occurred was suggested by the (nonsignificant) positive association found between amount of power attributed to the father in the mother-father and father-boy comparisons. However, other evidence indicates that this type of confounding did not occur extensively, if at all. First, about a dozen subjects were asked, "Who does he (she) say that to?" after the child responded to a mother-father comparison. In every instance, the child reported that the parent "said" the item to the other parent, not to an imaginatively introduced child. Second, although children perceived the father as significantly more powerful than the mother on the mother-father comparison, they did not attribute greater power to the father in the father-boy comparison than to the mother in the mother-girl comparison. Third, attribution of power to the father in the mother-father comparison was positively correlated with attribution of power to the mother in the mother-girl comparison ($p < .05$), a result opposite to that expected if confounding had occurred. It is also of interest that the mother-girl and father-boy comparisons were positively correlated ($p < .001$). It would appear that attribution of higher power to the father's sex role covaries positively with attribution of higher power to adult roles. Thus, rather than reflecting confounding, the pattern of correlations among the three measures indicates that children's use of the power dimension tends to generalize to both the sex- and age-role components of parent roles.

Discrimination of Child Sex Roles

Girls perceived the girl role as more positive (and less negative) than the boy role ($p < .001$), whereas boys did not consistently discriminate child sex roles by attitude. There were no significant variations with age for either sex. This finding is in essential agreement with that for children of 4½ to 6 years (7), although differences in the patterns of findings in the two studies lead to somewhat different interpretations. In the study of the younger sample, the fact that girls also discriminated adult sex roles on this dimension led to the conclusion that girls had developed sex-role concepts which generalize across the generations. This conclusion is not warranted in the present study, since girls did not discriminate adult sex roles by attitude. Perhaps this discrepancy between the studies reflects a true developmental change. However, for reasons discussed earlier, it would be important to replicate the findings on adult sex-role perception during the years 4 to 8 in order to evaluate this interpretation.

An interesting pattern of sex differences appears when this measure is broken down into its component items. As seen in Table 3, girls attributed more positive (and less negative) content to the girl role on all four items, particularly on items 2, 3, and 4. Boys, however, used the dimension significantly on fewer items, and not always in the same direction. The fact that both sexes attributed greater achievement competence (item 3) to themselves than to their opposite-sex peers is subject to two interpretations. It may signify a true sex difference in consensus on the definition of child sex

roles with respect to achievement competence. Or it may reflect the child's wish to attribute greater competence to himself than to another child, irrespective of the other child's sex. The latter explanation seems plausible in view of the achievement pressures placed upon middle-class children of both sexes. However, the same argument fails to account for the fact that boys did not perceive themselves as significantly more positive than girls on any of the other items. Indeed, on item 4, there was consensus within both sexes that girls are more likely than boys to do the "right thing" and to avoid trouble. What is clear from this pattern of findings is that both sexes use the attitude dimension to discriminate child sex roles, but, whereas girls apply it in the same way to a variety of settings, boys apply it more selectively and in a direction that varies among situations.

TABLE 3

SEX ATTRIBUTED POSITIVE CONTENT ON EACH ITEM FOR GIRL-BOY COMPARISONS

	SEX OF SUBJECT			
	GIRLS		BOYS	
Item Number	*Sex Attributed Positive Content*	*p*	*Sex Attributed Positive Content*	*p*
1	Girl	*ns*	Boy	*ns*
2	Girl	$< .01$	Girl	*ns*
3	Girl	$< .02$	Boy	$< .001$
4	Girl	$< .001$	Girl	$< .01$

The whole sample assigned more low power items to the boy's sex role than to the girl's sex role ($p < .02$), and this tendency did not vary significantly with the child's sex or age. This result is difficult to explain, as it contradicts rather well-established facts on sex differences in children's social behavior (24). It is also inconsistent with the finding on adult sex-role discrimination by power. The implausibility of this result raises a question about the validity of the power measure when applied to child sex-role discriminations. It will be recalled that low power items only were included in this measure because it seemed reasonable to use those items typically assigned to a particular generation when assessing discriminations within that generation. Perhaps this procedure resulted in loss of information that would have provided a more plausible finding, although the same procedure led to quite reasonable findings in the case of adult sex-role discriminations. Another possibility is that the power dimension collapses when the low end is applied to sex-role discriminations within the child generation. If this were the case, the measure could still consistently reflect a dimension, but one that differs from the intended dimension of power. Inspection of the

low power items suggests that this measure might have reflected self-centeredness, or simply differential social activity levels. It should be noted that the same question can also be raised about the dimensionality of the power measure applied to adult sex-role discriminations despite the fact that the findings on this measure were more plausible. Further study is required to resolve this problem.

Summary

This study investigated the development of children's conceptions of intrafamilial sex and age roles. The family role concepts of 225 middle-class children of 6 to 10 years of age were assessed by means of a modified paired-comparison procedure. The sample was divided into 12 subgroups based upon the child's sex and age. Intragroup and between groups analyses of role consensus revealed the extent of use of the dimensions of relative power (high vs. low) and attitudinal direction (positive vs. negative) to discriminate intrafamilial roles. The major findings were as follows:

1. In the discrimination of parent roles, there was much greater consensus on the power dimension than on the dimension of attitudinal direction.

2. Children discriminated age roles by assigning high power actions to the adult and low power actions to the child.

3. With increasing age, there was an increasing tendency to discriminate male but not female age roles by power.

4. Some age-role relationships were more differentiated by power than others. It was suggested that children age-type love-oriented more than object-oriented relationships.

5. The father's sex role was seen as more powerful than the mother's sex role.

6. There was a curvilinear relation between the child's age and the extent to which the father's sex role was perceived as more powerful than the mother's sex role.

7. Positive correlations were found among the measures of parent sex- and age-role discrimination by power, suggesting that use of the power dimension to discriminate parent roles generalizes across role situations.

8. Girls discriminated child sex roles by assigning positive actions to the girl and negative actions to the boy. In boys, the discrimination of child sex roles by attitude varied according to the specific interaction situation.

References

1. Asch, S. E. *Social psychology.* Prentice-Hall, 1952.

2. Baldwin, A. L., Kalhorn, J., & Breese, F. H. Patterns of parent behavior. *Psychol. Monogr.*, 1945, 58, No. 268.

3. Barker, R. G., & Wright, H. F. *Midwest and its children.* Row, Peterson, 1955.

4. Brim, O. G., Jr. The parent-child relation as a social system: I. Parent and child roles. *Child Develpm.,* 1957, 28, 343-364.

5. Bronfenbrenner, U. The study of identification through interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior.* Stanford Univer. Press, 1958. Pp. 110-130.

6. Emmerich, W. Parental identification in young children. *Genet. Psychol. Monogr.,* 1959, 60, 257-308.

7. Emmerich, W. Young children's discriminations of parent and child roles. *Child Develpm.,* 1959, 30, 403-419.

8. Gardner, L. P. An analysis of children's attitudes toward fathers. *J. genet. Psychol.,* 1947, 70, 3-28.

9. Gross, N., Mason, W. S., & McEachern, A. W. *Explorations in role analysis.* Wiley, 1958.

10. Hartley, R. E., Frank, L. K., & Goldenson, R. M. *Understanding children's play.* Columbia Univer. Press, 1952.

11. Heider, F. *The psychology of interpersonal relations.* Wiley, 1958.

12. Hoffman, M. L. Power assertion by the parent and its impact on the child. *Child Develpm.,* 1960, 31, 129-143.

13. Kagan, J. The child's perception of the parent. *J. abnorm. soc. Psychol.,* 1956, 53, 257-258.

14. Kagan, J., & Lemkin, J. The child's differential perception of parental attributes. *J. abnorm. soc. Psychol.,* 1960, 61, 440-447.

15. Mead, G. H. *Mind, self and society.* Univer. of Chicago Press, 1934.

16. Meltzer, H. Sex differences in children's attitudes to parents. *J. genet. Psychol.,* 1943, 62, 311-326.

17. Mussen, P., & Distler, L. Masculinity, identification, and father-son relationships. *J. abnorm. soc. Psychol.,* 1959, 59, 350-356.

18. Parsons, T. Family structure and the socialization of the child. In T. Parsons, R. F. Bales, & Associates, *Family, socialization, and interaction process.* Free Press, 1955. Pp. 35-131.

19. Radke, M. J. The relation of parental authority to children's behavior and attitudes. *Univer. Minn. Inst. Child Welf. Monogr.,* 1946, No. 22.

20. Sarbin, T. R. Role theory. In G. Lindzey (Ed.), *Handbook of social psychology. Vol. 1.* Addison-Wesley, 1954. Pp. 223-258.

21. Sears, R. R., Maccoby, E. E., & Levin, H. *Patterns of child rearing.* Row, Peterson, 1957.

22. SYMONDS, P. M. *The psychology of parent-child relationships.* Appleton-Century, 1939.

23. TASCH, R. J. The role of the father in the family, *J. exp. Educ.*, 1952, 20, 319-361.

24. TERMAN, L. M., & TYLER, L. E. Psychological sex differences. In L. Carmichael (Ed.), *Manual of child psychology.* Wiley, 1954. Pp. 1064-1114.

25. WARNER, W. L., MEEKER, M., & EELS, K. *Social class in America.* Science Research Associates, 1949.

26. WILSON, K. V. A distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1956, 53, 96-101.

CHILD'S SYMBOLIC CONCEPTUALIZATION OF PARENTS[1]

JEROME KAGAN,* BARBARA HOSKEN,
and SARA WATSON
Fels Research Institute

The language labels that the child applies to objects and events in his environment are an important determinant of much of his behavior. When the child learns to apply the word "dangerous" to objects with no feet that slither on the ground, his behavior toward this class of stimuli will be a function of the response that he has learned to the label "dangerous"—usually flight. This economical use of language to mediate a variety of behaviors is usually adaptive and allows the person to behave appropriately in a variety of new situations. However, the application of certain labels can lead to less desirable reactions. The semantic link between "disobedience" and "bad" often results in uncomfortable feelings when violation of parental standards has been committed or contemplated. Similarly, the semantic link "woman-flighty" predisposes an individual to react initially to all women as if this equation were universally true.

Parent figures are an important set of social objects with whom the child constantly interacts. The constellation of labels that the child ascribes to each parent will influence not only his reactions to them, but also his behavior with parent substitute figures (teachers, relatives, employers). Since the parents supply the child with his first definition of sex roles, the labels that are applied to father and mother are apt to generalize initially to all males and females.

During recent years there have been several investigations of the young child's conceptualization of parent roles (3, 4, 5, 7, 8). The results of these studies have been in remarkable agreement in indicating that the father, in relation to the mother, is perceived as more punitive, more feared, more

* Fels Research Institute, Yellow Springs, Ohio.

[1] This research was supported, in part, by research grant M-1260 from the National Institute of Mental Health, United States Public Health Service.

Child Develpm., 1961, 32, 625-636.

dominating, and less nurturant. This pattern of childhood associations linking father with aggressivity and mother with nurturance is consistent with the semantic differentiation of masculinity and femininity made by young adults (10).

It is generally assumed that the concepts *male* and *female* are linked with descriptive dimensions that are highly symbolic of these concepts. Psychoanalytic writings and anthropological reports suggest that dangerous animals, dark colors, and pointed objects are symbolic of masculinity; domestic animals, light hues, and rounded objects are presumably symbolic of femininity. In a recent empirical study with different cultural groups, it was found that angular lines, dark hues, and large objects were more frequently associated with men; their polar opposites were typically ascribed to women (9). It seemed profitable, therefore, to investigate the degree to which the young child had acquired these highly symbolic representations of sex roles.

The primary purpose of this research was to investigate the child's differential conceptualization of the concepts *mother, father,* and *self* on a variety of symbolic dimensions that are presumably linked both with aggressivity and nurturance, as well as masculinity and femininity. An ancillary goal was to replicate the results of earlier studies (7, 8) that indicated a perception of the father as more feared and more potent than the mother.

Method

Subjects and Procedure

Ss were 98 white children (59 boys and 39 girls) from intact families. Their ages ranged from 6 to 8 years with a modal and median age of 6½ years. All *Ss* were attending the first or second grade in one of three elementary schools at the time of data collection. Each *S* was individually interviewed by one of two female experimenters (BH or SW) on three separate occasions. The sessions were separated by a period of 7 to 14 days with a modal intersession interval of 10 days. Several days prior to the first test session the *E* visited the child's classroom in order to become acquainted with each *S* and reduce the effect of stranger anxiety on the child's test performance.

During the first test session, half of the *Ss* were questioned about their conceptualization of their father; the remaining half were tested for their conception of mother. During the second session each child was questioned about the second parent. On the third session three series of questions were administered: (a) the child's conception of himself, (b) his conception of the concept "bad," and (c) a picture completion test.

The Tests

Symbolic conceptualization of father, mother, self. The technique used to assess *S*'s conceptualization of *father, mother,* and *self* consisted of 66

pairs of picture stimuli. Each pair illustrated the polar opposite aspects of a descriptive dimension. The 11 dimensions sampled were: strong-weak, big-little, nurturant-nonnurturant, competent-incompetent, punitive-nonpunitive, dangerous-harmless, dirty-clean, dark hue - light hue, cold-warm, mean-nice, and angular-rounded. Each of these 11 dimensions was presented six times with six different pairs of relevant pictures. The pictures contained no illustrations of men or women nor any clue that the child might interpret as directly suggestive of a male or a female.

After *E* had gained the rapport of the child, she said, "I'm going to show you some pictures, and I want you to tell me which one reminds you of your (father, mother, self). Which one is like your (father, mother, self)?"

E then showed the *S* some practice stimuli that were obvious in their appropriateness for men, women, or children (e.g., trousers vs. dress, apron vs. men's shoes, child's toy vs. piece of paper). *E* then administered the 66 test stimuli. *E* read a verbal description of the two pictures on each card that contained the salient symbolic dimension. *E* would say, for example, "Here is a strong rabbit, and here is a weak rabbit. Which one reminds you of your (father, mother, self)?" or, "Here is a mouse that knows how to read, and here is a mouse that does not know how to read. Which one reminds you of your (father, mother, self)?"

The presentation of the 11 dimensions was randomized as was the position of the pictures in each pair. *E*'s verbal descriptions of the 66 pairs of pictures follow:

1. Strong rabbit and weak rabbit.
2. Little boat and big boat.
3. Someone giving money to child and someone not giving anything.
4. Mouse that knows how to read and mouse that does not know how to read.
5. Someone who is mad at child and someone who is not mad at child.
6. A lamb and a lion.
7. A clean dog and a dirty dog.
8. Black horse and a white horse.
9. Cold bowl of soup and a warm bowl of soup.
10. Mean whale and a nice whale.
11. Round shape and a pointed shape.
12. Weak piece of wood and strong piece of wood.
13. Big ball and little ball.
14. Someone who does not want to pet the dog and someone who likes to pet the dog.
15. Bird who is lost and doesn't know where to go and bird who does know where to go.
16. Someone not slapping child's hand and someone who is slapping child's hand.
17. A tiger and a horse.
18. Dirty foot and a clean foot.
19. White bear and black bear.
20. Warm fireplace and cold fireplace.
21. Nice cat and a mean cat.
22. This shape and that shape.

23. Strong fence that stood up and weak fence that fell down.
24. Little table and big table.
25. Someone covering the cat so it will be warm and someone not covering the cat.
26. Dog that can do tricks and a dog that can not do tricks.
27. Someone telling child to go to (his) (her) room and someone not doing this.
28. A hammer and a gun.
29. Clean room and a dirty room.
30. Brown table and a red table.
31. Cold day and a warm day.
32. Mean squirrel and a nice squirrel.
33. This line and that line.
34. Strong plate that did not break and weak plate that did break.
35. Big block and a little block.
36. Someone not kissing the baby and someone kissing the baby.
37. Squirrel that doesn't know where to look for food and a squirrel that does know where to look.
38. Someone not yelling at child and someone yelling at child.
39. Crocodile and a bird.
40. Dirty pig and a clean pig.
41. White telephone and black telephone.
42. Warm sweater and a cold sweater.
43. Nice cow and a mean cow.
44. This design and that design.
45. Strong rope and a weak rope.
46. Little telephone and big telephone.
47. Someone bringing the lamb into the house and someone telling the lamb to go away.
48. Pitcher that will pour the right way and pitcher that won't.
49. Someone scolding the child and someone not scolding the child.
50. A rabbit and a snake.
51. Clean ash tray and a dirty ash tray.
52. Purple hat and a yellow hat.
53. Cold bed with thin blanket and a warm bed with a thick blanket.
54. Mean elephant and a nice elephant.
55. This line and that line.
56. Weak cloth that tore and strong cloth that did not tear.
57. Big cup and little cup.
58. Someone who will not help calf out of the wire and someone who will help the calf get loose.
59. Clock that doesn't remember to ring and a clock that does remember.
60. Hand that will not spank the child and a hand that will spank the child.
61. A knife and a ruler.
62. Dirty bed and a clean bed.
63. Light blue pencil and dark blue pencil.
64. Warm bath and cold bath.
65. Nice mouse and a mean mouse.
66. This shape and that shape.

This set of 66 stimuli was administered on each of the three sessions to assess the child's conception of father, mother, and self. After *E*'s verbal description of each stimulus, *S* pointed to the picture he felt reminded him of his father, mother, or self.

The concept of "bad." On the third test session *S* first went through the 66 stimuli indicating which picture best applied to himself. *E* then asked the child if he knew the meaning of the word "bad." When *E* was satisfied that the child was familiar with the word she said, "Now we are going to look at these pictures again and I want you to tell me which one is bad. Which picture is the bad one?" *E* then administered the first 44 stimuli.

Picture completion: Perception of parents. The final test in the third session consisted of 12 pictures (separate pictures for boys and girls) in which a boy or girl was illustrated in a different situation but no adult was illustrated. *E* said, "Here are some pictures which show a (boy, girl) but the mommy or daddy is missing. You tell me who is missing, the mommy or the daddy."

The pictures were designed to suggest parental nurturance, punitiveness, or strength. *E*'s verbal description for each of the 12 pictures follows:

PICTURE 1. Someone is giving the (boy) (girl) an ice cream cone. Who is it, the mommy or the daddy?

PICTURE 2. Someone is going to get the (boy) (girl) the toy (he) (she) can't reach. Who is getting it, the mommy or the daddy?

PICTURE 3. Someone just spanked the (boy) (girl). Who spanked (him) (her), the mommy or the daddy?

PICTURE 4. Someone is picking the (boy) (girl) up in the air. Who is doing it, the mommy or the daddy?

PICTURE 5. Someone is yelling at the (boy) (girl). Who is yelling, the mommy or the daddy?

PICTURE 6. Someone is giving the (boy) (girl) a present. Who is giving the present, the mommy or the daddy?

PICTURE 7. The dog is chasing the (boy) (girl) and (he) (she) is running to someone for help. Who is the (boy) (girl) running to, the mommy or the daddy?

PICTURE 8. The (boy) (girl) is reading a book with someone. Who is (he) (she) reading with, the mommy or the daddy?

PICTURE 9. Someone is going to spank the (boy) (girl). Who is it, the mommy or the daddy?

PICTURE 10. Someone is taking a walk with the (boy) (girl). Who is it, the mommy or the daddy?

PICTURE 11. Someone is on the seesaw with the (boy) (girl). Who is it, the mommy or the daddy?

PICTURE 12. The (boy) (girl) is mad at someone. Who is (he) (she) mad at, the mommy or the daddy?

RESULTS

Symbolic Conceptualization of Parents and Self

For convenience of presentation only one pole of each of the 11 dimensions will be considered. Since *S* had to choose between the two poles, the scores for either pole accurately reflect the nature of the results. The 11

attributes to be discussed are strong, big, nurturant, competent, punitive, dangerous, dirty, dark, cold, mean, and angular.

The data were initially analyzed for differences with respect to experimenter, school, and order of questioning (father or mother in first session). None of these differences even approached statistical significance, and the data for experimenters, schools, and order of questioning were pooled in the presentation to follow. Table 1 presents the average score on the 11 dimensions for father, mother, and self for the sexes separately and for the total group.

TABLE 1

AVERAGE SCORE ON CONCEPTUALIZATION OF FATHER, MOTHER, AND SELF

	FATHER			MOTHER			SELF		
Dimension	*Boys*	*Girls*	*Total*	*Boys*	*Girls*	*Total*	*Boys*	*Girls*	*Total*
Strong	5.6	5.9	5.7	4.7	5.1	4.9	5.6	4.8	5.4
Large	5.7	5.8	5.7	4.1	4.1	4.1	4.9	3.5	4.4
Nurturant	5.2	5.2	5.2	5.0	5.6	5.3	5.2	5.5	5.3
Competent	5.4	5.2	5.3	5.2	5.3	5.2	5.4	5.2	5.3
Punitive	0.9	0.9	0.9	0.8	0.4	0.6	0.7	0.6	0.7
Dangerous	1.8	1.7	1.8	0.6	0.4	0.5	1.8	0.4	1.3
Dirty	0.6	0.7	0.7	0.2	0.1	0.2	0.7	0.5	0.6
Dark	3.0	3.0	3.0	2.0	1.8	1.9	2.7	1.7	2.3
Cold	0.6	0.5	0.6	0.8	0.3	0.6	1.1	0.8	1.0
Mean	0.8	0.9	0.8	0.5	0.1	0.4	0.8	0.6	0.6
Angular	2.7	2.0	2.4	1.9	1.3	1.6	2.2	1.2	1.8

Differential conception of parents. Boys and girls agreed that father, in relation to mother, was stronger, larger, more dangerous, more dirty, darker, and more angular. Each of these differences between maternal and paternal attributes was significant at the .01 level or better (all p values are for two tails). Girls described the father as more punitive and meaner than the mother ($p < .05$), but differences on these two dimensions were not significant for boys. There were no significant differences between the parents for the attributes nurturant, cold, or competent.

Conception of self. Boys, in comparison to girls, labeled themselves as stronger, larger, more dangerous, darker, and more angular. Each of these differences was significant at the .01 level or better. These five dimensions were also more frequently ascribed to the father than the mother by both sexes. Thus, boys and girls conceptualized themselves as more similar to the same than the opposite sex parent on the dimensions strong, large, dangerous, dark, and angular.

In order to obtain a more direct measure of the similarity between the child's conceptualization of himself and the same sex parent, difference

scores were computed for each dimension between the score for each parent and the score for the self. For example, the number of times the child labeled himself as "strong" was subtracted from the number of times he labeled his father or mother as "strong." These differences between the scores for the self and each parent were computed for all 11 dimensions. The mean difference between father and self was larger for girls than for boys ($p < .05$); the mean difference between mother and self was larger for boys than for girls ($p < .05$). Thus, the 6-year-old child labels himself as more similar to the same than to the opposite sex parent on attributes that are highly symbolic of the sex roles.

TABLE 2

AVERAGE "BAD" SCORE FOR EACH DIMENSION

Dimension	*Boys*	*Girls*	*Total*
Strong	0.8	0.6	0.7
Large	1.5	1.3	1.4
Nurturant	0.3	0.1	0.2
Competent	0.5	0.5	0.5
Punitive	3.8	3.8	3.8
Dangerous	3.9	4.0	3.9
Dirty	3.8	3.8	3.8
Dark	2.7	2.7	2.7
Cold	3.3	3.6	3.4
Mean	4.0	3.9	3.9
Angular	2.9	3.0	2.9

Conception of the Word "Bad"

Table 2 presents the average number of times a dimension was called "bad" by the child (range 0 to 4).

Boys and girls were again in close agreement that the dimensions, mean, dangerous, dirty, punitive, cold, weak, nonnurturant, and incompetent, were "bad" (all had average scores over 3.0). The high degree of conceptual agreement between boys and girls is reflected in the rank-order correlation of +.99 for the differential assignment of the word "bad" to the 11 dimensions.

Picture Completion Test for Perception of Parents

Table 3 presents the percentages of boys and girls responding mother or father to each of the 12 pictures.

These data agree with those obtained with the more symbolic stimuli. On pictures 3 and 9, which illustrate physical punishment of the child, boys and girls agreed that father was more likely to be punitive than

TABLE 3

PERCENTAGE OF GROUP RESPONDING MOTHER OR FATHER TO THE TWELVE PICTURES

Picture	Boys		Girls		Total	
	Father	*Mother*	*Father*	*Mother*	*Father*	*Mother*
1. Child receiving ice cream	54.4	43.8	26.3	71.0	43.2	54.7
2. Person reaching for toy	63.2	36.8	55.3	42.1	60.0	38.9
3. Child spanked	63.2	36.8	68.4	31.6	65.3	34.7
4. Child picked up	84.2	15.8	94.7	5.3	88.4	11.6
5. Child yelled at	43.8	54.4	50.0	50.0	46.3	52.6
6. Child receiving present	22.8	75.4	23.7	68.4	23.2	72.6
7. Child fleeing dog	66.7	31.6	52.6	47.4	61.0	37.9
8. Child read to	42.1	57.9	23.7	76.3	34.7	65.3
9. Child to be spanked	71.9	28.1	71.0	28.9	71.6	28.4
10. Child walking	33.3	66.7	28.9	71.0	31.6	68.4
11. Child on seesaw	70.2	22.8	28.9	57.9	53.7	36.8
12. Child mad	50.9	45.6	71.0	26.3	58.9	37.9

mother ($p < .01$). On pictures 2, 4, and 7, which illustrate scenes requiring strength, size, and the ability to protect the child from danger, boys and girls both answered father more than mother ($p < .05$). To pictures 6 and 10, which illustrate nurturance and companionship, *S*s responded mother more than father ($p < .01$). Responses to these pictures indicate that strength and punitiveness were viewed as more characteristic of the father while nurturance was ascribed to the mother.

Answers to some of the pictures yielded sex differences. To pictures 1, 8, and 11, which suggest nurturance, girls were more likely to reply mother than the boys ($p < .02$, .10, and .01 respectively). To picture 12 (child mad at someone) girls were more likely to answer father than were boys ($p < .05$). Thus, girls were more likely than boys to view the mother in a nurturant light and the father in a hostile one. This pattern agrees with the girls' responses to the 11 symbolic dimensions. Girls labeled the father as more punitive and meaner than the mother; boys did not make this distinction between the parents.

DISCUSSION

Most of these results are in essential agreement with previous findings that were based on different samples of children and different procedures (3, 4, 5, 7, 8). The child views the father as more punitive, more fear arousing, more potent, but less nurturant than the mother.

Responses to the symbolic stimuli did not indicate that nurturance was more characteristic of the mother. This result seems due, in part, to the procedure used with this instrument. In previous research, the child was presented with a description of a nurturant act or object and *forced to choose between mother or father* as the more appropriate parent for the stimulus. Under these conditions the child typically chose the mother. In the procedure used in this study, the child was not forced to choose between mother or father and was allowed to label both parents as nurturant. This is exactly what these *S*s did. The stimuli used to illustrate the nurturant dimension forced the child to pick between a nurturant event and one characterized by an absence of nurturance. All of the children agreed that lack of nurturance was a "bad" characteristic, and they were reluctant to assign a "bad" trait to either parent. The picture completion test, which used the more popular procedure of forcing the child to pick between mother or father, yielded results that were in agreement with previous investigations. The mother was labeled as more nurturant than the father on pictures 6, 8, and 10.

The present data verify some of the classical hypotheses regarding symbolic manifestations of sex roles. Angularity of line, dark hues, dangerous animals, and large objects appear to be symbolic of the father and of masculinity in the conceptual organization of the school age child. The work of Jenkins and Russell (6) and Osgood (9, 10) indicate the similarity in conceptualization of sex roles between children and adults. When college students rated the concepts of father and mother on the semantic differential, the father received higher scores than the mother on the dimensions, cruel, successful, strong, and bad (6). A recent study (9) of Anglo, Japanese, Navajo, and Mexican-Spanish adults used a procedure similar to the one used in this research and yielded similar results. *S*s were presented with 13 pairs of visual stimuli that signified dimensions such as verticality, thickness, angularity, size, and darkness. *S* matched one of these pairs of stimuli with concepts such as man, woman, bad, heavy, good, fast, happy, and strong. The results were in essential agreement with those found with the present group of children. For example, the dark stimulus was labeled "bad" and was more often associated with "man" than with "woman" by all four language-culture groups. Similarly, the rounded line design, as opposed to the angular design, was more frequently assigned to the concept "woman" than to the concept "man." Largeness, on the other hand, was assigned more often to "man" than to "woman." Thus, there appears to be a universality surrounding the dimensions angular, large, and dark that applies to children and adults as well as to people from different language cultures. It seems fair to conclude that some of the labels that are implicitly linked with the concepts male and female in the adult psyche have been acquired by the time the child is 5 or 6 years old.

The language learning that accounts for this communality in the conceptualization of sex roles probably occurs very early, perhaps as soon as

language becomes available to the child. The degree of homogeneity between child and adult and across language cultures suggests that factors outside the child's immediate family are partly responsible for this language learning. All fathers are not more punitive or more dangerous than mothers, yet most children respond as if this were so. It is necessary to turn to the consequences of certain universal experiences as well as to the effect of the mass media in order to explain this phenomenon.

In western culture the major sources of communication portray the male role as markedly more aggressive and violent than the female role, and most adults concur in this evaluation (1, 2, 6). It is reasonable to expect, therefore, that animals and objects that are usually linked with aggression and danger will be associated with masculinity. At a point further out on the generalization continuum, it is plausible to assume that any stimulus dimension that is associated with pain, danger, or anxiety will be preferentially assigned to the male role. Some of our children verbalized that the "angular" designs can "hurt you," and the association between pointed objects (needles, nails, knives) and pain may account for the association matrix linking angularity with masculinity. The association between "darkness" and "father" could be a function of the more frequent use of dark clothes by men as well as the association cluster that includes darkness, the night, and danger.

It is possible that this generalized perception of men as more dangerous and stronger than women influences each child's perception of his own father. That is, once the child has acquired the association between danger and men and nurturance and women, he may begin to anticipate aggression from his father and affection from his mother. The content of these sex role concepts can also affect an individual's initial posture with strange males and females. If child and adult agree that boys and men are stronger, larger, and more dangerous than girls and women, an initial attitude of timidity and passivity will be more likely to prevail with strange men than with strange women.

It is interesting to note that on both the symbolic stimuli and the incomplete pictures, girls labeled the father as more hostile and punitive than did boys. This same finding emerged with an independent sample of 67 children who averaged one year younger than the present sample (8). The replication of this result reinforces the conclusion that the young girl views the father with more fear than the young boy does. It is suggested that, because the 6-year-old boy has begun to identify with his father, he is forced to repress some of the father's fear arousing qualities.

The Evaluative Dimension

The attributes, strong, large, nurturant, competent, clean, warm, nice, and nonpunitive, were most frequently assigned to both parents. The opposite poles of weak, small, nonnurturant, incompetent, dirty, cold, mean, and punitive were infrequently assigned to either parent. These latter attributes were labeled "bad" by the child. It appears that the child's reluctance

to assign "bad" characteristics to his parents was a major determinant of his answers to the 66 symbolic stimuli. Rank-order correlations between the average "bad" score for each dimension and its frequency of assignment to the parents were —.75 and —.94 for father and mother, respectively. Thus, the assignment of traits to the parents was regulated, in large measure, by the child's evaluation of the negative qualities of that dimension.

This pervasive influence of a good-bad dichotomy on the child's responses agrees with Osgood's conclusion that the evaluative dimension is one of the primary structuring principles that people use in classifying people, objects, and events (10).

Summary

A group of 59 boys and 39 girls, age 6 to 8, were interviewed on three separate occasions for their conceptualization of their parents and themselves. A special instrument was devised to assess the child's conceptualization of father, mother, and self on 11 polar dimensions: strong-weak, big-small, nurturant-nonnurturant, competent-incompetent, punitive-nonpunitive, dangerous-harmless, dirty-clean, dark-light, cold-warm, mean-nice, and angular-rounded. In addition, the child was asked to supply the missing figure (father or mother) to each of 12 scenes illustrating children in different situations.

Boys and girls conceptualized the father, in comparison to the mother, as stronger, larger, darker, more dirty, more angular, and more dangerous. The children conceptualized themselves as being more similar to the same sex than to the opposite sex parent. Answers to the incomplete pictures revealed that the mother was labeled as more nurturant but less punitive and less potent than the father.

It was suggested that the child's contact with certain common, environmental contingencies and communications from the public media were important factors in the symbolic association between masculinity and the dimensions of angularity, darkness, size, strength, and danger. The semantic domains of the concepts male and female held by the 6-year-old overlap in large measure with the conceptualizations of the adult.

References

1. Bennett, E. M., & Cohen, L. R. Men and women: personality patterns and contrasts. *Genet. Psychol. Monogr.*, 1959, 60, 101-153.
2. Child, I. L., Potter, E. H., & Levine, E. M. Children's textbooks and personality development: an exploration in the social psychology of education. *Psychol. Monogr.*, 1946, 60, No. 279.
3. Emmerich, W. Young children's discriminations of parent and child roles. *Child Develpm.*, 1959, 30, 403-419.

4. FINCH, H. M. Young children's concept of parent roles. *J. Home Econ.*, 1955, 47, 99-103.

5. GARDNER, P. L. An analysis of children's attitudes toward fathers. *J. genet. Psychol.*, 1947, 70, 3-28.

6. JENKINS, J. J., & RUSSELL, W. A. An atlas of semantic profiles for 360 words. *Amer. J. Psychol.*, 1958, 71, 688-699.

7. KAGAN, J. The child's perception of the parent. *J. abnorm. soc. Psychol.*, 1956, 53, 257-258.

8. KAGAN, J., & LEMKIN, J. The child's differential perception of parental attributes. *J. abnorm. soc. Psychol.*, 1960, 61, 440-447.

9. OSGOOD, C. E. The cross-cultural generality of visual-verbal synesthetic tendencies. *Behav. Sci.*, 1960, 5, 146-169.

10. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning.* Univer. of Illinois Press, 1957.

ASSESSING SOME ASPECTS OF THE PARENT-CHILD RELATIONSHIP

F. N. Cox* *and* Patricia M. Leaper
University of Melbourne

In this study we wished to make a systematic examination of some selected aspects of parent behavior on the adjustment of a diverse, yet typical group of Australian primary school children. In selecting the aspects of parents' behavior we would study, we decided to rely largely on propositions formulated by psychoanalytically-oriented clinicians and anthropologists. We felt it would be an advantage, however, to assume that certain components of syndrome complexes of parent behavior can be isolated and studied separately. It was thought that such an assumption would simplify the process of evolving operational definitions of the aspects of the parent-child relationship we chose to study.

Theoretical Considerations

The first problem, then, was that of choosing which of the many aspects of parent behavior we would examine. As has been implied, selection of variables was strongly influenced by the heuristic value of psychoanalytically-oriented propositions. Nearly all writers with this point of view have emphasized the supreme importance of "love" for child adjustment. "Restriction" is also frequently stated to have significant effects, while "family cohesion" or "family harmony" is another factor which some of these writers claim can be important for understanding children's adjustment.

Love

In our thinking about this variable we were particularly influenced by the writings of Anna Freud (7, 8), Erikson (6), Jersild (11), and Balint (3). From certain of their propositions, we tried to evolve a definition of "love" which would serve as a basis for constructing measuring instruments. More especially, we attempted to derive a definition which would delineate a universe of behaviors which constitute "love," and to do this in such a

* Department of Psychology, University of Melbourne, Melbourne, Victoria, Australia.

Child Develpm., 1961, 32, 637-649.

way that we could then construct items to sample this universe.

The generalized definition of "love" adopted in this study is that the parents' behavior should be expressive of a set of attitudes which we term child oriented, developmentally relevant, and child respecting. The *child oriented* attitude is expressed in using childrearing methods that take account of what the child actually needs, as well as what the parents think he needs, and what they might want him to need. An important implication is that parents do not repeatedly use their child as a means of gaining narcissistic satisfactions.

The *developmentally relevant* attitude is expressed by parents setting expectations and demands that are appropriate to the emotional, physical, and social capacities and skills of the child at a given stage of his development. In order to ensure that the child will not be consistently forced to cope with insoluble problems, parents will set definite and relevant goals for him. This implies that parents will have actively familiarized themselves with what can be reasonably expected of children at various ages.

The *child respecting* attitude is expressed by parents showing appreciation of the child's individuality. They recognize and endorse the child's emotional, intellectual, and physical strengths and limitations, including his susceptibility to urgent, internal stimulation. They also grant their child some autonomy and privacy.

It is apparent that this type of definition excludes a number of notions concerning "love" that psychoanalytic and other writers have stressed. For instance, it is not a definition which emphasizes demonstrativeness or warmth; it does stress, however, the importance of well-informed guidance. Moreover, this definition differs in detail from the propositions of any single writer in this field, but we thought that it would delimit a universe of behaviors from which a set of items for assessing variations in parent behaviors could be derived. The nature of the items would vary from one age level to another, and items at any age level would have to be related to the known facts of development at that age.

Restriction

The importance of control, limit-setting, and overprotection have been discussed by such writers as Baldwin (1), Erikson (6), Anna Freud (7), Jersild (11), Levy (12), Lewin (13), and Redl and Wineman (15). We attempted to derive a definition of "restriction" which covered as many of the propositions of the above workers as could be synthesized into a coherent pattern. That is, we tried to delineate a unitary but broad universe of parental behaviors, ranging from neglect through control (limit-setting) to overprotection. While this approach may not lend itself to covering all of the extremes of overindulgence and overrestriction described by Lewin and Levy, it does have the advantage of making an a priori distinction between "restriction" and the above propositions referring to the guidance aspects of "love." We defined our universe of "restriction" as consisting in

those parental behaviors which confine children's freedom of action. It was argued that degrees of restriction could be measured by assessing the extent to which parents demand that their child participate in household routines, confine his activities to specified locations, and associate only with designated persons. While failure to supervise any of the child's activities probably implies neglect, overrestriction is implied when parents fail to ensure that their child has some time to himself, chances to explore his neighborhood, and opportunities for mixing with people living around him.

Family Cohesion

Various terms, such as atmosphere, integration, cohesion, and harmony, have been used to describe the family as a group. The variety of these terms indicates that the group functions of a family can be looked at from different points of view. It is apparent that the different points of view arise because of different predictive purposes.

For instance, it seemed to us that the propositions advanced by Lewin (14) and Baldwin (2) differ from those advanced by writers such as Erikson (6) and Friedlander (9), in that Lewin and Baldwin have not been concerned with differentiating symbiotic from other family relationships. For the limited purposes of the present study, we decided to define family cohesion in terms of the degree to which the various functions necessary for the maintenance of the family unit are apportioned according to mutually agreed preferences and skills. We argued that these family functions consist of providing and managing finance, food, and shelter; initiating and consolidating the social relationships of the family unit to the extended family, the neighborhood, and the wider community; and providing facilities and opportunities for constructive use of leisure time. We decided to use measures of mother, father, and child joint participations in the functions of initiating and consolidating social relationships and constructive use of leisure time as indices of family cohesion, for we intended to use children's reports on family activities. It seemed improbable that children would be able to give information about the nature of the division of sustenance functions within the family. Further, even in those cases in which children are capable of observing the division of sustenance functions they may misinterpret them, particularly if the division is so well established that discussion about it rarely focuses upon central issues.

Family cohesiveness will be expressed when the parents join with each other to visit and entertain, enjoy casual leisure activities in the home, and go out on family excursions and include their children whenever their activities are appropriate.

The variables of love, restriction, and cohesion comprise those aspects of the parent-child relationship we chose to study in detail. Many people will argue that other significant factors could be profitably substituted for our variables. In particular, many will claim that "consistency" of parental actions is of paramount importance for predicting child adjustment. How-

ever, it appears to us that consistency can be regarded in two ways: either as a modifier of other variables or as a factor having independent effects on child behavior. We did not feel at all confident that the methodology to be described below would be appropriate to a serious study of consistency and its effects, for the study relies on children's accounts of activities in the home.

Methodology

Basic Assumptions

The methodology adopted in this study rested upon two interrelated assumptions: namely (a) that 10- to 11-year-old children can give as accurate an account of certain specified aspects of parent behavior as parents can and (b) that children of this age are sufficiently mature to report what usually happens without merely reflecting the events of a particular day.

Population

To initiate our studies of whether the variables of Love, Restriction, and Family Cohesion could be measured by administering structured question schedules to children, we decided to begin with a restricted population. The population specifications were that (a) the children should be boys in the fifth grade at primary (elementary) school and, therefore, the great majority would be 10 or 11 years of age; (b) the sample should be drawn from a wide cross-section of the local urban community; and (c) the children should, if possible, come from intact family units and be living at home.[1]

These specifications were set down to ensure that the children were capable of writing answers to clearly worded questions;[2] that changes in development associated with puberty would not complicate the initial attempts to measure the family variables; and that the extrafamilial influences to which the children were exposed were as uniform as possible—for example, the broad course of their school activities would be approximately equivalent. Further, we wanted to concentrate on the construction of measures which would, in the first instance, distinguish parental behaviors in "normal," sociologically intact family units.[3]

Secondary Assumptions

The basic assumption that the specified population of children will give accurate information about parental behavior will only obtain if the informa-

---

[1] These studies were carried out in the city of Melbourne, the population of which is approximately 1,750,000.

[2] Mentally defective children and migrant children who cannot understand and write English were deliberately excluded.

[3] It is apparent that this specification is designed to exclude institutionalized children and children who do not live with both parents.

tion is sought under the following conditions: (a) if questions relevant to a given variable sample a wide variety of the situations in which that variable is likely to operate; (b) if children have had frequent and repeated opportunities to observe the parental behaviors about which they are questioned; (c) if questions are directed toward eliciting statements of fact rather than of opinion; (d) if questions are framed in such a way that they do not arouse undue emotional reactions. Emotionally provocative questions would have an unknown effect upon the veracity of answers to the question itself and to those that follow it, quite apart from stimulating unnecessary and possibly uncontrollable anxieties in the children.

It was soon apparent that no existing questionnaire appropriate to the specified population was designed so that these conditions obtained. However, perusal of existing questionnaires suggested a particularly useful technique for the construction of question schedules—the daily life schedule. It seemed that this technique would simplify presentation of provocative questions, and so help to gain children's acceptance of, and cooperation in, the task. Further, it was thought that the use of the daily life schedule technique would facilitate the adequate sampling of situations in which the Love, Restriction, and Family Cohesion variables would be likely to operate. Presentation of separate items in a time sequence also has the effect of dispersing items relevant to a particular variable. Finally, the technique enables interpolation of "buffer items" with face validity, which in itself would seem to help to divert attention from, and reduce reaction to, some necessarily emotionally potent items.

Form and Content of Scales

All of the Love items consisted of questions about situations in which sympathetic and discreet parental participation is necessary for the optimal psychological well-being of a 10- to 11-year-old child. Specifically, children were asked if parents exercise firm but sympathetic control over the ways in which they, the children, begin and end their day; if, and how, parents provide opportunities for communication about the children's reactions to pleasurable and unpleasurable experiences; and about the quality of parental reactions to especially good and bad child behaviors.

The form of the items intended to measure Love varied. Most consisted of a related series of open-ended and closed questions, the former requiring only short answers. For example, the form of an item about the quality of parental reactions to especially good child behaviors was: (a) "If you do something especially good, does someone in your family praise you for doing it?" (b) "Put a ring or circle around the person or people who praise you." M F B S b s O.[4] (c) "How do these people praise you?" Scoring of such items was necessarily interpretative and depended upon the patterns

[4] The symbols M, F, B, S, b, s, O designated mother, father, elder brother, elder sister, younger brother, younger sister, and other adults living with the family, respectively. These symbols were used in many items and the children had no difficulty with them.

of answers to the series of questions comprising one item. The remainder of Love items consisted of single multiple-choice, closed questions.

Specific items in the Restriction scale elicited information about the extent to which a child participates in household routines, has to confiine his activities to specific locations, and has to restrict his associations with other people. Restriction items were similar in form to the Love items. Scoring of responses to Restriction items was interpretative.

The Family Cohesion items consisted in simple, multiple-choice questions about whether parents join with each other to visit and entertain, enjoy casual leisure activities in the home, and go out on family excursions.

In addition to including buffer items in the question schedules, we asked children some questions about the sociological characteristics of their families, including parental occupations, size of family unit, age of family members, and variety of family possessions.

The preliminary form of the question schedule was divided into two parts. It was found that, when the 74 items were arranged in a time sequence, they fell into a natural division of questions about weekday and weekend activities. We decided to use this natural division to provide a point at which administration could be broken.[5]

Preliminary Study

The two preliminary forms were administered to a group of 45 fifth grade children in May 1956.

Administration

Because of the length and complexity of question schedules, it was found necessary to use experienced proctors. Proctors were used to ensure that children gave individual and meaningful answers to each question. Special features of the administration included reading each item aloud twice; emphasizing that children were to give the answer which described what *usually* happened; and presenting the schedule as a survey of the range of activities in which children of the local community engage. These and other aspects of the administration may be illustrated by the following excerpts from the standardized instructions given to the children before the Ordinary Day question schedule. "We are trying to find out the ways in which children spend their days. . . . The questions start with your getting up in the morning and end with your going to bed at night. We want you to give the answer which is *usually true for you.*"

Administration of the preliminary forms took two and a half hours. In addition, a 20-minute play break was given between administration of the

[5] The two question schedules were called "My Ordinary Day" and "My Usual Weekend."

Ordinary Day and Usual Weekend. The children enjoyed the idea of telling about their everyday activities and seemed to find most questions interesting. Form and phrasing of questions was on the whole appropriate: children gave satisfactory answers to all but a few items. Further, none of the children seemed disturbed by the content of the questions.

Scoring

Criteria for scoring the Love, Restriction, and Family Cohesion items were derived from a priori definitions of these variables. For example, criteria for scoring a Love item about the quality of parental reactions to especially good child behaviors were that parental interest be shown by the giving of praise; that interest be shown by mother *or* father or *both*;[6] and that praise be expressed by giving appropriate rewards. Similarly, in scoring a Restriction item about whether children are expected to confine their activities to specified locations, we first established that the answer indicated the existence of a restriction, but then ascertained whether the restriction was a defensible one. For example, we did not score an item as an indicator of Restriction if places like swiftly running creeks at which many fatalities occur were forbidden to children. Since Family Cohesion items were in multiple choice form, scoring was largely mechanical.

Findings

An internal consistency analysis[7] showed that 40 of the 74 items could be arranged into four clusters.

Love. It was found that only three of the eight items designed to measure this variable were strongly related to each other.

Restriction. The items intended to measure this variable divided into two clusters. The first cluster consisted in six questions about the extent to which children are permitted to undertake activities and social relationships of their own choosing and to choose the venue of some of these. We decided to call this cluster of items *Social Restriction.* The remaining nine questions were about whether children assist their parents in routine household activities and/or take responsibility for these. This cluster was named *Household Responsibilities and Duties.* Since these two item clusters were unrelated, it seemed that our a priori assumption that excessive participation in household routines must restrict children is probably untenable.

Family Cohesion. This consisted of a cluster of 22 items. It was found that in more than 80 per cent of cases in which parents are reported to share activities, children state that they, too, participate in these activities.

In all cases, the clusters of items which comprised the various measures of our variables were found to be independent of each other and of all other

[6] Initially, no attempt was made to discriminate between the behavior of fathers and mothers.

[7] This analysis was undertaken on the Hammond Matrix Sorter.

items in the question schedules. All items comprising a cluster were significantly correlated with the total cluster score.

Second Study

The findings from the preliminary study encouraged us to continue our attempts to describe selected aspects of the parent-child relationship in terms of children's reports of their daily lives with their parents. In the second study, we aimed, first, to cross-validate the findings concerning the independence and internal consistency of the clusters of items representing the Social Restriction, Household Responsibilities, and Family Cohesion variables on a larger and more representative sample of children. Further, we wanted to continue the attempt to construct a set of items to measure Love.

The Ordinary Day and Usual Weekend schedules were reduced in size by rejection of nondiscriminating items. Sociological and buffer questions were retained. Five of the eight items designed to measure Love were reworded, and scoring criteria sharpened. In all, the revised form of the question schedules contained 59 questions, 45 of which were directed at the measurement of Love, Social Restriction, Household Responsibilities and Duties, and Family Cohesion variables.

The revised schedules were administered to a second sample of 150 fifth grade boys in July 1956. This sample was drawn from four schools, two of which were in residential suburbs, while the other two were in or adjacent to industrial areas. The administrative procedures were unchanged, but the time taken to administer both revised question schedules was reduced to an hour and a half.

An analysis of the eight items intended to measure Love showed that six of these formed a cohesive cluster. This cluster consisted of the three items which were found to be satisfactory in the preliminary study and three of the five revised items. A further examination of the six items comprising this cluster showed that the average intercorrelation of the items was of the order of .81. This very high degree of internal consistency was obtained from items asking children about the extent to which their parents' behavior was oriented to their needs, respecting of their personality characteristics, and designed to guide them in their relationships with their environment.

In analyzing the relationships between the four item clusters, the sample of 150 boys was divided randomly into two groups. Since we had examined intact school grades, we could not assert that this sample was typical of children of the age group being studied: for instance, it was possible that one of the schools examined had an undue proportion of children from a homogeneous socioeconomic background. In both samples, it was found that the item clusters which constituted the measures of the four main variables were independent of each other.

TABLE 1

CORRELATION COEFFICIENTS BETWEEN FAMILY VARIABLES

	Love	*Social Restriction*	*Household Duties*	*Family Cohesion*
Sample 1 ($N=75$)				
Love		—.17	.06	.01
Social Restriction			—.06	—.11
Household Duties				—.01
Family Cohesion				
Sample 2 ($N=74$)				
Love		.07	.02	—.06
Social Restriction			.08	—.12
Household Duties				.05
Family Cohesion				

NOTE.—The samples are unequal because one case was rejected.

The findings summarized in Table 1 replicate the independence of the Social Restriction, Household Responsibilities and Duties, and Family Cohesion dimensions. This table also shows that the Love item cluster is independent of the other three.

The earlier finding that, within a given cluster, all items were related to the total score at at least the .05 level of significance was replicated in this second study. In addition, it was found that the scoring was capable of being reproduced by an individual unfamiliar with the question schedules. The obtained index of interscorer reliability was 97 per cent.

THIRD STUDY

In this study we planned to investigate some correlates of scores on the four family scales if the finding that the scales were independent of each other was cross-validated.

The third sample was selected to meet the same general specifications as those set out for earlier samples: namely, only literate 10- to 11-year-old boys, living in intact families in the local urban community would be included. In this study, we made a serious attempt to ensure that the sample drawn should be large and that children from different sections of the community should be represented. However, it was impractical to examine members of one large parochial group. The sample may be considered to be representative of all other sections of the local community. In all, 519 boys attending 16 different schools were examined in February and March 1957. The schools were located in a wide variety of ecological areas: at least two schools were drawn from each of inner industrial, inner residential,

outer industrial, outer residential and mixed areas.[8] Children from both government and nongovernment schools were included in the sample.

A number of measures of potential correlates of family variables were administered to the third sample. Some of the measures of the correlates were administered at the same time as the family question schedules, while the remaining correlate measures were administered two to three months later. For the purposes of investigating relations between the family scales and the correlate measures, we specified that, for a case to be included in the analysis, data had to be complete. If even one correlate measure was incomplete or invalid, the case was dropped. Unfortunately, a severe epidemic of Asian "flu" started in March 1957. This led both to delay in the administration of the second set of measures and to an extraordinarily high rate of absence from school, and the latter factor led to a large number of incomplete cases. In all, 432 complete case records were available for analysis.

TABLE 2

CORRELATION COEFFICIENTS BETWEEN FAMILY VARIABLES ($N = 432$)

	Love	*Social Restriction*	*Household Duties*	*Family Cohesion*
Love		—.08	.09	.03
Social Restriction			—.08	—.13
Household Duties				.12
Family Cohesion				

Analysis of the characteristics of the family scales showed that, for the third sample, the Love, Social Restriction, Household Duties, and Family Cohesion scales were relatively independent of each other. Relevant correlation coefficients are presented in Table 2.

Reliability of the Scales

Because of the relatively interpretative nature of the scoring, we rechecked interscorer reliability. Twenty-five cases were drawn at random from the original pool of 519 cases. Two persons who had not previously worked on records independently scored all the items constituting the family scales in these 25 records. The agreement between the two independent scorers was 99 per cent; moreover, each of these scorers registered 99 per cent agreement with the original scorers.

An opportunity arose to adminster the question schedules to 46 children who had participated in the second study. The time interval was eight

[8] The local council of the community in which the study was undertaken lays down regulations concerning the type and extent of industrialization permitted in a given area.

TABLE 3

RETEST RELIABILITY COEFFICIENTS FOR FAMILY SCALES ($N = 46$)

Scale	*r*
Household Duties	.78
Social Restriction	.71
Family Cohesion	.60
Love	.80

months. As the results in Table 3 show, all of the reliability coefficients are significant beyond the .01 level.

In addition to calculating these simple retest reliability coefficients, we examined the retest consistency of the single items. On the Household Duties scale each single item was consistent beyond the .001 level, despite the eight-month interval. Five out of the six items comprising the Love scale were consistent beyond the .001 level. While the remaining Love item was not significantly consistent, scores tended to indicate that the item was not wholly unreliable. In the case of Social Restriction, all six items were consistent beyond the .001 level. Only six out of the 12 items[9] constituting the Family Cohesion scale were significantly consistent.

Validity

While the data already presented show that it is possible to use children's statements to evolve measures of defined aspects of the parent-child relationship which are independent, internally consistent, and reasonably stable, no evidence has yet been presented to indicate that the measures are related to independent assessments of child or parent behavior.

A number of techniques for assessing child behavior and personality were administered to children comprising the sample used in the third study. An indication of the predictive utility of the family scales can be given by setting out some relationships found to obtain between the family scales and some of the independent measures of child behavior. A significant and meaningful relationship, $r = -.39$, between the Love scale and Reputed Immaturity (4) was obtained for a sample of 250 used in the third study. No other family scale was related to this sociometric measure. The Household Duties and Responsibilities scale was found to bear a linear relation with the external criterion of school marks. For a sample of 96 boys, drawn from four schools where marks had a comparable mean and standard deviation, $\chi^2 = 6.47$ with 2 df ($.05 > p > .02$). School marks were not related to scores on any other family scale. Neither the Family

[9] The Family Cohesion scale was reduced from 22 to 12 items after the second study.

Cohesion nor the Social Restriction scales have so far shown any consistent relations with independent measures of child behavior.

It was also possible to relate scores on the family scales to data on parent and child behavior obtained during interviews with parents. Mothers of 69 of the children were interviewed by an experienced psychologist or social worker who knew no details of the work already done with the children. There were eight interviewers, no one of whom did less than six interviews. Although interviewers used their own chosen techniques of phrasing and ordering questions, they were provided with a schedule of topics. The interviews covered a wide range of topics including socioeconomic status of family, childrearing practices and beliefs, and parents' aspirations and plans for their children. Topics were included which gave mothers an opportunity to discuss their behavior and the rationale for their behavior in situations we thought sensitive to the expression of "Love" and "Household Duties and Responsibilities" variables. Mothers were also questioned about their children's actions and attitudes in these situations. The interviewers, who were not acquainted with the children's scores on the family scales, were asked to record interview data in a way that would allow the mothers' responses to be clearly differentiated from interviewer interpretations. The present writers jointly scored the mothers' recorded statements in terms of the a priori definitions of Love and Household Duties and Responsibilities. The scoring was blind, all identifying particulars having been removed from the interview records beforehand.

The interviews took place between 18 months and two years after administration of the family scales to the children. Nonetheless, on both the Love and the Household Duties and Responsibilities variables, scores from interview records were significantly related to scores from question schedules administered to children. On Love the correlation coefficient was .29, and on Household Duties and Responsibilities the correlation coefficient was .39.[10] When records obtained by different interviewers are examined separately, the relation between interview scores and child question schedule scores is seen to vary considerably: for instance, the correlation coefficients for Love range from .85 to —.09.

In view of these findings, it seems reasonable to conclude that the two scales measuring Love and Household Duties and Responsibilities are psychologically meaningful as well as internally consistent and independent of each other. Despite the internal consistency and independence of the Social Restriction and Family Cohesion scales, we cannot present unequivocal evidence of the validity of these scales.

More generally, the results presented in this article indicate that, if care is taken in the selection and wording of questions, 10- to 11-year-old children can give reliable and valid information about defined aspects of the parent-child relationship. While no claim is made for universal validity of

[10] For the sample of 69 cases, the Love correlation coefficient is significant beyond the .05 level, and the Household Duties and Responsibilities beyond the .01 level.

the variables isolated and measured in this study, results obtained do seem to encourage the view that construction of measuring devices can profitably follow from a priori definition of variables.

References

1. Baldwin, A. L. The effect of home environment on nursery school behavior. *Child Develpm.*, 1949, 20, 49-62.
2. Baldwin, A. L., Kalhorn, J., & Breese, F. H. Patterns of parent behavior. *Psychol. Monogr.*, 1945, 58, No. 268.
3. Balint, M. *Problems of human pleasure and behaviour.* London: Hogarth, 1956.
4. Beswick, D. G., & Cox, F. N. Reputed aggression and dependence in children. *Austral. J. Psychol.*, 1958, 10, 144-150.
5. Child, I. L. Socialization. In G. Lindzey (Ed.), *Handbook of social psychology.* Vol. 2. Addison-Wesley, 1954. Pp. 655-692.
6. Erikson, E. H. *Childhood and society.* Norton, 1950.
7. Freud, A. *The ego and mechanisms of defence.* London: Hogarth, 1937.
8. Freud, A., & Burlingham, D. *Infants without families.* London: Allen & Unwin, 1943.
9. Friedlander, K. *The psycho-analytical approach to juvenile delinquency.* London: Routledge & Kegan Paul, 1947.
10. Herbst, P. G. Analysis and measurement of a situation: the child in the family. *Hum. Rel.*, 1953, 6, 113-140.
11. Jersild, A. *Child psychology.* (4th Ed.) Prentice-Hall, 1954.
12. Levy, D. M. *Maternal overprotection.* Columbia Univer. Press, 1943.
13. Lewin, K. *A dynamic theory of personality.* McGraw-Hill, 1935.
14. Lewin, K., Lippitt, R., & White, R. Pattern of aggressive behavior in experimentally-created "social climates." *J. soc. Psychol.*, 1939, 10, 271-299.
15. Redl, F., & Wineman, D. *Children who hate.* Free Press, 1951.

ADAPTABILITY: A PROBLEM SOLVING APPROACH TO ASSESSING CHILDREARING PRACTICES[1]

IRVING TALLMAN*
San Francisco State College

Theories concerning the consequences of parental practices on children's behavior have only recently been interpreted in terms of empirical observations. Present research in this field can be broadly categorized under three theoretical orientations: those concerned with specific parental behavioral dimensions which serve as antecedents to consequent personality and behavioral components in children (3, 20, 24); those which assess parental attitudes and the consequences of such attitudes (19); and, finally, those which attempt to assess interaction patterns and the differentiation of roles within the family (14, 17, 21).

This paper will report on the foundational work and research findings based on still another theoretical approach. In this approach emphasis is placed on the problem solving character of the childrearing process. The data to be reported are based on partial results of a study presently in progress at San Francisco State College. One phase of this study is concerned with the relationship of parents to their retarded children. What follows is not an advocacy of the superiority of a particular approach over other means of assessing parent-child relationships and practices. Rather, the purpose of this paper is to present a unique approach and one which harbors the potential for contributing an additional perspective to the understanding of the parent-child relationship.

A PROBLEM SOLVING FRAMEWORK

The conceptual framework to be presented has been developed principally out of the theoretical formulations of John Dewey, H. G. Mead, and W. I. Thomas. The basic assumption made is that the parent-child relationship exists as a series of problem solving situations. Such situations involve

* San Francisco State College, 1600 Holloway Avenue, San Francisco 27, California.

[1] The conceptual and empirical work reported in this paper was made possible by a grant from the United States Office of Education, Department of Health, Education, and Welfare, Contract No. SAE 6439. The writer wishes to acknowledge his indebtedness to Samuel Levine for his careful reading and criticism of the manuscript, as well as for his helpful comments and suggestions over the three years in which the ideas to be presented in this paper were formulated.

Child Develpm., 1961, 32, 651-668.

mutual expectations as to the behavior of each participant. From the point of view of the parent, he not only has expectations for the child in a situation but also of himself in terms of his parental role. His actions are seen as affecting or training the child to behave according to certain standards. A parent's actions can be conceived as having two purposes: the resolution of immediate problems and the socialization of the child.

A parent responding to a child in any situation would theoretically have an infinite number of behavioral acts to draw upon. That this is not the case is due to a number of factors which can best be summarized under the dual categories of personal experience and social structure. On the basis of past experience the parent has accepted some behavioral forms and discarded others. His life experiences have positively reinforced some forms of action and negatively reinforced others. The social structure, by means of normative standards, also delimits the choice of acts available. The normative pattern prescribes the rights and duties of a particular social position, and in so doing, structures the forms of interaction possible in any situation.

Within the limits of these experiential and social restrictions on behavior, we must assume that a wide range of choices of action still remain. No society could long exist if it did not allow sufficient behavioral flexibility to cope with constant changes and unexpected situations. From a problem solving point of view, the normative patterns can provide only the skeletal structure for effectively dealing with a situation. It follows that no *specific action* or form of behavior can be considered as right or best for any individual: first of all, there exists a range of possible "right" responses; secondly, the responses available to the individual are limited by experience and the normative patterns of the culture and the subculture. For example, a middle class mother may respond to a child's misbehavior by giving him "love" or reassurance, whereas a working class mother may spank her child under the same circumstances. Either action, if adhered to rigidly and mechanically might be fruitless and unsuccessful whereas, if it were appropriate to the situation, it could be effective. To consider the former as loving behavior and the latter as punitive would overlook the more crucial meaning of the behavior, e.g., its appropriateness to the situation.

The problem of raising a child is a day-to-day and moment-to-moment affair, and, within certain structural limits, it occurs in constantly changing and shifting situations. The action a parent chooses, therefore, is determined by his perception of the specific situation. It is the "definition of the situation" which W. I. Thomas has referred to as the essential stimulus for action (22). Interactions and transactions which occur in a situation effect constant changes and call for constant situational redefinitions.

This dynamic interpretation of behavior presents a conceptual problem to the researcher who must make his assessment at a particular point in time. The theoretical framework presented in this paper suggests at least a partial solution to the problem, that is, the recapitulation of the problem solving process. It is assumed that in facing any situation an individual will

bring into play his experiences with similar situations ar.d his awareness of alternative forms of action. These two factors in combination with the person's assessment of the situation will result in a plan and finally specific action. This process, hereafter referred to as adaptability, will be amenable to empirical exploration to the extent that it can be verbalized and observed.

The following pages will be concerned with explicating the concept of adaptability, defining its components, providing a set of operational definitions, and, finally, reporting data relevant to the utility of the concept to research in childrearing.

The Concept of Adaptability

Baldwin (2) states that the term ability has been defined in two senses: it reflects what an individual "can" do, and it can also be used as a primary factor basic to a specific act. In either case it is considered as ". . . a necessary but not sufficient condition for actual performance" (p. 199). A distinction is therefore made between being able to accomplish a feat and actually doing so. Baldwin further states that ". . . ability is a meaningful concept only because it describes the characteristics of a person which contribute to his carrying out adaptive behavior" (2, p. 210). Adaptive behavior is "result oriented"; it is ". . . behavior oriented to the consequences of action." Thus, a problem solving analysis demands more than a specification of ability; it must include basic conditions under which adaptive behavior occurs. In childrearing situations we are interested in the appropriate abilities, *plus* other factors, which contribute to the parent's behaving *adaptively* toward his child. Baldwin suggests that one of the added requirements necessary to convert the ability into adaptive action is motivation. The writer would add to this the concept of environmental opportunity or environmental difficulty, for, as Baldwin recognizes, some situations may demand greater skill than others for achieving a similar result.

Four components are conceived of as essential to effective adaptive behavior toward a child. Two of these may be considered as abilities, the other two as factors which make it possible to employ the abilities effectively. These latter elements are motivation and the availability of resources in the environment. The two attributes which can be considered as abilities are empathy and flexibility. These four components combine to form what Burgess and Wallin (5) have titled "adaptability." The term aptly combines the two fundamental constructs which the writer considers as necessary for effective problem solving.

In order for behavior to be effective the four components of adaptability are conceived of as acting in combination. Motivation is seen as the basic prerequisite. If the parent does not wish to do anything about a situation, we assume that nothing will be done, whatever his potential problem solving skills. Conversely, the parent may be highly motivated, but if he does not have the understanding and flexibility necessary to determine the appro-

priate action and respond accordingly, the probability is that the action taken will not be adaptive. In addition, neither empathy nor flexibility should produce effective results without benefit of the other. The ability to understand the child and the situation (empathy), while lacking a repertory of alternative actions (flexibility), appears as fruitless as having a wide repertory of actions without such action being guided by appropriate understanding.

Only three of the four components are employed in the actual determination of adaptability. "Availability of resources" is difficult to assess and is, at least in part, incorporated in the construct of flexibility.

The research instrument developed, therefore, was designed to measure independently the components of motivation, empathy, and flexibility. In this way it was possible, indirectly, to test the hypothesis that these components are interdependent. The findings relevant to these interrelations will be discussed in a subsequent section of this paper.

The Problem of Definition

Terms such as motivation, empathy, and flexibility are encumbered with a great degree of ambiguity. The problem is to define each of the components with sufficient specification to meet the purposes of our particular conception of adaptability, while, at the same time, not destroying the constitutive meaning which does exist for these terms.

Parent adaptability is defined as the parent's ability to deal effectively with problematic situations by changing roles, attitudes, and actions in terms of new or modified understandings of the situation with which he is confronted.[2] With this definition in mind we will turn to an examination of each of the components.

Flexibility is defined in both attitudinal and behavioral terms. It represents an attitude which individualizes the child in the light of a specific situation, as opposed to stereotyping both the child's behavior and the situation. In addition, it represents the ability to employ alternative means of action toward one's child in the light of the parent's evaluation of the situation, as opposed to acting in a rigid, stereotyped manner.[3] This continuum involves more than the use of alternative modes of actions in deal-

---

[2] This definition of adaptability is a modification of that presented by Burgess and Wallin (5). These writers define adaptability as "the capacity of the person to change his roles, his attitudes, and his behavior in order to adjust to those of other persons or to new or modified situations" (p. 623).

[3] The definitions of six levels on the continua for each of the components of Flexibility, Empathy, and Motivation and the interview schedule have been deposited as Document number 6838 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C. A copy may be secured by citing the Document number and by remitting $2.50 for photoprints, or $1.75 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

ing with a situation; it also includes the evaluation of these actions and the integration of attitudes with stated actions. The lower end of the flexibility continuum is characterized by absolutism and arbitrariness; the higher responses indicate an ability to evaluate relative and pertinent influences. To this extent the scheme approximates Rokeach's concept of the "open and closed mind" (16). In addition, low flexibility contains attributes which contribute to what Frenkel-Brunswick has termed "intolerance of ambiguity" (8). Therefore, a response rated as low flexibility would be one which stereotypes the child in the situation and employs stereotypic statements to rationalize the action or lack of action taken. Thus, a statement such as "retarded children need to be loved, so what we do is love him" would be rated low. A middle response would be a single behavioral response to the specific situation, for example, "If he interrupts me when I talk with my husband I'll say, 'I'm talking now, wait until I finish.'" A high response is one which chooses a specific behavior from a number of alternatives, the behavioral choice being dependent on the parent's specific evaluation of the situation, for example, "I find if you show him how to do it by taking his hand and doing it with him, from then on he knows what to do. It's much easier than telling him because it's hard for him to understand."

Empathy has variously been defined as "a gift of understanding people" (1), "social sensitivity" (10), and "the ability to take the role of the other," an ability which Sarbin (18) subsumes under what he calls an "as if" skill. Our interpretation of empathy adheres to these general definitions. It differs, however, from the usage given to it by Dymond (7), Notcutt and Sylva (14), and others in that we make no stipulation as to the accuracy of perception. An attempt to assess perceptual accuracy would involve delimiting the situational cues a parent may have to just those areas in which the investigator has factual information. Such a procedure would be a digression from a problem solving orientation. In addition, as Bruner and Taguiri (4) have pointed out, perception of emotions requires as many cues as possible, as well as an understanding of the situational context in which the perceived behavior occurs. Our interest is in the depth and inclusiveness of the parent's perception of the child, within the context of a situation. The assumption is that inclusiveness and the integration of pertinent situational factors in understanding the child is sufficient evidence of sensitivity to infer the ability to take the child's role and consequently to justify the use of the term empathy.

Empathy is defined, for our purposes, as the parent's ability to interpret his child's behavior in terms of feelings, thoughts, and motives and to relate this understanding to the specific situation in which the parent and the child are involved. The lowest point of the continuum represents the inability of the parent to think or express why the child behaves as he does in a given situation. The midpoint is indicated by the parent's making a single or simple cause-and-effect relationship between the child's activity or response and the stimulus which elicited the response; the highest level

is understanding the child's feelings, thoughts, and motivations and combining this understanding with the stimulus situation in order to explain the behavior.

The term motivation has perhaps the widest range of interpretation and the least amount of constitutive meaning of any of the components. Its breadth and complexity is dramatically summarized by Murphy's statement that "Any serious conception of social motivation must be theoretically complex enough to provide raw material for the manifold subtle components in the motive potential of mankind, must see how the growth and learning process may complicate and interweave these expressions, and how the different types of cultural arrangements may mold, stabilize, and integrate the various fusions and complications of individual experience" (12, p. 613).

The magnitude of the task, as outlined by Murphy, is overwhelming and goes beyond the pragmatic requirements for assessing a component in adaptability. The statement illustrates the theoretical and communicative problems resulting from employing the idea of motivation. What is necessary is to specify the *kind* of motivation that is of interest. We are not concerned with basic or genetic drives; our focus lies in a particular form of derived social motive, specifically the motivation of the parent to solve interpersonal problems in which he and his child are interacting. A consequence of such motivation is involvement. Involvement represents the degree to which the parent gives of his time and energy in order to achieve certain results. Motivation, for parent adaptability purposes, is therefore defined as an affectual investment characterized by involvement of the parent in effecting certain anticipated goals for the child. "Affectual investment" is inferred from the behavioral interpretation of involvement. Involvement is conceived as a continuum, its lowest level being an unwillingness on the part of the parent to recognize any responsibility for the child or the situation, i.e., "There is nothing you can do with him." At the other end of the continuum, high involvement includes evidences of increased effort, such as inclusion of external resources and engaging in persistent follow-through by means of interaction with the child, i.e., "I went to school and watched the teacher work with him. She's calm, but she's firm and she expects him to sit quietly and eat, so I expect better behavior from him at the dinner table. If he doesn't behave, he must leave the table. He'll object, but I insist." The middle of the continuum is represented by a willingness on the part of the parent to stay with the situation over a period of time, i.e., "I always talk to her and warn her against going out in the street."

Underlying the concept of adaptability, and consequently each of the components, are two continua. The first conceives of the parent's perception of the situation running from generality to specificity. In line with the conception of behavior as being determined by the definition of the situation, the assumption is that, the greater the ability to specify the elements in the specific situation, the greater the potential for problem solving.

TABLE I.

RELATION OF THE GENERALITY-SPECIFICITY AND COGNITION CONTINUA TO EACH OF THE ADAPTABILITY COMPONENTS

	ADAPTABILITY COMPONENTS			UNDERLYING CONTINUA FOR ALL COMPONENTS	
	Motivation	*Empathy*	*Flexibility*	*Generality-Specificity*	*Cognition**
Low	No acknowledgement of responsibility for personal involvement.	No recognition or thoughts about causation of child's behavior.	Rigid and stereotyped attitudes and actions.	Categorical or general response.	"Unguided behavior."
Middle	Staying with a situation by a particular type of action.	Making a simple and single connection between cause and effect of children's behavior.	Use of a specific or limited act or applied formula.	Limited specificity.	"Simple guided behavior."
High	Integration of parents' efforts with outside resources; interpersonal interaction with child with persistent follow-through.	Understanding of the child's feelings, motivations, and thoughts as well as the situational factors which resulted in behavior.	Dealing with the situation in terms of its specific meaning, while employing past experience and recognizing changes in the child.	High specificity.	"Cognitive guidance."

* Based on Baldwin (2).

The other underlying continuum evolves from Baldwin's concept of adaptive behavior as behavior oriented to the consequences of action (2). If this is the case, then the more the behavior is what Baldwin has termed "cognitively guided," the greater is the probability of effective problem solving. The continuum is conceived as moving from behavior which would remain unvaried under any circumstances (i.e., unguided); through a midpoint in which the behavior is guided (i.e., a simple reaction to the situational stimuli); to behavior which is cognitively guided (i.e., consciously employed to effect specific and long range goals).

As Table 1 indicates, these underlying continua are common to the components of flexibility, empathy, and motivation. They are a priori determinations which are integral to the concept of adaptability.

Operational Interpretation and Instrument Development

It can be seen, from the above discussion, that the conceptual framework of adaptability places great emphasis on the conscious thought processes of the parent. The assumption is that people who can think in adaptable terms are more inclined to act in a problem solving manner. If we make the additional assumption that people who are able to express themselves in adaptability terms are more likely to act in this way, we can think of employing an interview schedule as our research instrument. The test of this assumption lies in the experimental use of the instrument. Empirical support of the postulates to be tested can be considered as indirect support of the instrument and its assumptions. However, if the postulates are not supported, what are we to say about the assumptions and the instrument? Will the results be due to an inappropriate instrument, or have our theoretical constructs been disproved? The writer, at this point, cannot argue for the proven utility of the research instrument. He can, however, point out that in his use of such an instrument the results supported the hypothesis that the children of highly adaptable parents showed greater growth on a measure of social competence than children of lower adaptable parents.

An interview was devised to elicit parental responses on each of the three adaptability components. The questions were open ended and worded in a simple and direct manner. The respondent was presented with a problematic parent-child situation, such as behavior at the dinner table, and then asked specific questions designed to educe flexibility, empathy, and motivation responses.[4] Every attempt was made to make the interview as natural as possible while at the same time maintaining a standardized and structured procedure. The interviewer introduced a situation as follows: "I'd like to talk with you now about Johnny's behavior at the dinner table." After asking a general question about the parent's impression of Johnny's table manners, the flexibility question was asked in this way: "How do you go about train-

[4] *See* Footnote 3.

ing Johnny in table manners?" If necessary, it was followed by a further probe, "Is this what you always do?" Then, the empathy response was elicited by asking, "Why do you think Johnny behaves (as parent has previously stated) at the dinner table?" The motivation questions were somewhat more complex. First, an attempt was made to obtain the parent's level of aspiration for the child in the given situation, e.g., "What would you hope for (or like) in regard to Johnny's behavior at the dinner table?" The parent was then asked two questions designed to determine the amount of involvement evident in achieving these goals for the child. The next question was, "What do you think can be done to help achieve this?" and, if necessary, "What have you done in this regard?"

This, in general, was the procedure employed with the parents. Interviewer bias was, of course, one of the principal sources of error in this type of approach. In an attempt to control as much of this error as possible, interviewers were given fairly extensive training in the use of the schedule. The schedule was adhered to rather rigidly, and interviewers were allowed to probe only under specified conditions. Interviewers were all professional social workers, clinical psychologists, or educational psychologists.

In the present study verbatim transcriptions were made from the tape recorded interviews, and the subject's responses for every question were entered on separate slips of paper. The slips were stamped with a random number representing the subject. Responses were then grouped according to the question number and component. Judges would rate 50 such responses at a time. Judges were paired, each making an independent rating for every response and then meeting with his partner to discuss and resolve any disagreements. No pair of judges rated all responses for any group of subjects. This method would seem virtually to eliminate halo effect. In all, judges rated 10,000 separate responses. The reliability for independent ratings of pairs of judges is high (correlations were consistently above .90). Judges rated responses according to the criteria established for each of the components along 6-point continua.

In addition to the above procedure, judges made ratings on the basis of listening to the tape recorded interviews for a random sample of 30 parents. The correlations between the two rating procedures was .84. This suggests that the rather cumbersome and expensive method employed to avoid halo effect may be unnecessary.

Scoring Procedures

The method for obtaining an individual's empathy, flexibility, or motivation score derives from the procedure employed in developing the instruments and, in particular, the method of selecting the questions. This selection procedure was, in turn, related to certain assumptions relevant to the conception of adaptability.

Questions were selected on the basis of 60 pre-experimental interviews. An original pool of 43 situations were compiled, reflecting the opinions of

staff members as to common yet problematic areas of parent - retarded child interaction. In addition, the literature was reviewed, and other scales and schedules were examined for further cues as to relevant situations.

Five revisions of the interview schedule were made, based on the 60 pre-experimental interviews. The final schedule included 24 different situations.

There were three principal criteria for the inclusion of a question: (a) the degree to which the question elicited responses pertinent to our constructs of flexibility, empathy, and motivation; (b) the frequency with which parents indicated the situation occurred; (c) the range of responses the question pulled which differentiated among respondents. It will be noted that no attempt was made to order the questions; they merely represent a sample of situations found to be relevant to parents of retarded children. The response to *each* question is scored on the basis of an a priori ordered 6-point scale, based on the theoretical constructs relevant to the particular component being measured. However, inasmuch as the questions themselves cannot be ordered as to meaningfulness, difficulty, intensity, etc., the procedure for arriving at an individual score must be classificatory (23). In other words, this procedure does not allow us to meet the axioms necessary for an ordinal or interval scale; and, therefore, if we were to sum the questions scores in order to obtain a component score, we would be doing violence to the logic of measurement.

The decision to use the above procedure of item selection, rather than developing a formal scale along the lines of Guttman or Thurstone, stems from the theoretical assumptions regarding the nature of abilities. We assume that individuals seek to resolve problems as quickly and economically as possible. Therefore, even the most adaptable individual will respond only to the degree necessary to achieve his ends. Thus, if a situation does not represent a serious problem, the parent will not be affectually aroused and therefore not overly involved. As a consequence he will probably respond in a routine manner. The same circumstance will hold true if the situation has been resolved in the past and no longer presents a problem.

It is only on the questions which pose situations that are perceived as problematic by the parent that we hope to differentiate between low and high adaptable parents. The high parent, it is predicted, will respond to the challenge by an increase in cognitively guided problem solving activities; the low adaptable person will respond by increased rigidity and stereotyping. The former will react progressively; the latter will respond regressively.

It can be seen from the above discussion that there is no means for predicting which situational questions will be perceived as problems for any respondent. What is claimed is that the questions, having been adequately pretested, represent a sufficient sample of common problematic situations so that a proportion of them will be pertinent to almost every parent.

The scoring procedure assumes that at least a majority of the questions will evoke problematic experiences in any parent. If there are 15 questions

for a component, a highly adaptable parent should score high on at least five or more of the questions for that component. Conversely, a low adaptable parent will score low on five or more of the questions. The probability is that the parent who is given an over-all high rating for a component will score no lower than 3 (middle) on a given question; nor will a parent who received a low component rating score higher than 3 on a given question. The scoring procedure may be summarized as follows: (a) Judges rate parent's responses to each question on a 6-point scale. (b) Parent's ratings are then examined and the frequencies of 1 and 2, 3, 4, and 5 and 6 responses computed. If the subject has five or more questions scored at the 1 and 2 levels, he is ranked low; if five or more ratings are at the 5 and 6 levels, he is ranked high. If the subject has five or more level 4 responses he is given an upper middle ranking. All other distributions are given a middle ranking.

The validity or utility of this categorical system must rest on its ability to predict to an external criterion. Unfortunately, there is no exact method of determining "internal validity." [5]

Scores for each of the three components are arbitrarily combined by summation. This procedure is employed principally for the sake of convenience and in lieu of more precise knowledge of how these factors combine. However, there does appear to be sufficient evidence to support combining the components in some form. Table 2 presents the correlations for each of the questions with the components and with total adaptability.[6]

It appears that each of the items contributes sufficiently, not only to its component, but also to total adaptability. There are, however, a few questions which do not appear to be contributing either to component scores or to adaptability (such questions would, of course, be eliminated in any subsequent use of the instrument). In addition, the instrument seems more consistent for mothers than for fathers.

Table 3 reports the correlation of each of the components with total adaptability. Again, it can be seen that, although each component is operating independently, there is a significant association with total adaptability. In addition, the three components produce significant associations with an external criterion only when they are combined.

---

[5] Measures of internal consistency such as odd-even reliability and item total score correlations (*see* Table 2) were sufficiently high to support the conventional claim for a homogeneous scale. (The odd-even Pearson correlation, using the Spearman-Brown correction for attenuation, was .84 for mothers and .81 for fathers.) As noted above, the writer does not claim the adaptability instrument to be a scale in the technical sense; these correlations do suggest, however, that the situational questions are appropriate and meaningful to most parents.

[6] All correlations reported are gamma coefficients, unless stated otherwise. Gamma is a nonparametric statistic which measures the directed order of association, the degree to which the order of x is generally the same as the order of y. Specifically:

$$\text{Gamma} = \frac{\text{probability of like orders minus the probability of unlike orders}}{\text{sum of like and unlike probabilities}}$$

TABLE 2

ITEM CORRELATIONS (GAMMA) WITH SUBSCALES AND TOTAL SCORE FOR MOTHER AND FATHER SAMPLES

	Mother ($N=113$)				Father ($N=100$)			
*Item**	*Flex.*	*Emp.*	*Mot.*	*Adapt.*	*Flex.*	*Emp.*	*Mot.*	*Adapt.*
Flexibility Items								
2.	.53	.42	.36	.51	.53	.27	.21	.42
3.	.52	.37	.40	.51	.41	.48	.26	.30
4.	.46	.29	.28	.41	.55	.26	.29	.42
5.	.37	.24	.25	.33	.19	.03	.09	.09
7.	.58	.26	.34	.45	.55	.29	.19	.38
8.	.59	.32	.38	.50	.50	.34	.21	.47
9.	.36	.24	.20	.30	.48	.21	.14	.32
10.	.63	.45	.38	.54	.48	.25	.22	.36
11.	.48	.38	.37	.48	.37	.09	.25	.24
13.	.55	.32	.31	.45	.52	.27	.35	.46
16.	.61	.31	.40	.50	.25	.16	.16	.22
17.	.38	.27	.13	.30	.40	.05	.13	.22
18.	.41	.23	.11	.31	.30	.20	.39	.34
19.	.27	.04	.08	.16	.33	.23	.30	.33
20.	.42	.34	.34	.42	.48	.06	.19	.26
Empathy Items								
2.	.32	.15	.21	.27	.26	—.004	.20	.18
3.	.21	.18	.17	.21	.28	.27	.23	.32
4.	.26	.06	.16	.20	.32	.02	.08	.16
5.	.46	.22	.25	.35	.35	—.06	—.05	.00
7.	.34	.28	.29	.36	.39	.13	.18	.30
9.	.40	.19	.16	.29	.46	.21	.27	.40
10.	.34	.05	.31	.24	.28	—.05	.07	.16
11.	.48	.24	.38	.42	.39	.23	.04	.31
12.	.42	.39	.28	.42	.42	.23	.23	.39
13.	.46	.36	.40	.46	.36	.50	.20	.47
16.	.46	.28	.20	.35	.49	.36	.26	.52
17.	.47	.27	.23	.38	.40	.10	.02	.27
18.	.46	.35	.18	.37	.55	.37	.24	.54
19.	.43	.22	.34	.38	.56	.23	.24	.49
20.	.49	.29	.37	.45	.47	.30	.30	.47
Motivation Items								
6.	.31	.20	.32	.32	.45	.30	.18	.43
9.	.44	.30	.28	.39	.58	.25	.10	.42
11.	.50	.32	.34	.48	.42	.12	.14	.26
14.	.35	.22	.26	.31	.41	.25	.32	.41
15.	.37	.34	.30	.40	.34	.21	.03	.26
17.	.53	.16	.36	.39	.43	.26	.22	.39
20.	.35	.30	.24	.35	.49	.25	.25	.41
21.	.43	.32	.35	.44	.48	.30	.20	.47
22.	.48	.22	.20	.36	.43	.24	.31	.37
23.	.48	.33	.36	.44	.34	.25	.08	.29
24.	.44	.24	.25	.38	.33	.17	.18	.24
25.	.51	.25	.23	.40	.41	.45	.28	.37
26.	.16	.27	.42	.20	.37	.11	.23	.27

* Item refers to question on Parent Adaptability Schedule.

TABLE 3

ADAPTABILITY SUBSCALE INTERCORRELATIONS (GAMMA)

	Motivation	*Flexibility*	*Adaptability*
Mothers ($N=113$)			
Empathy	.48	.49	.72
Motivation	..	.48	.75
Flexibility	..	..	.75
Fathers ($N=100$)			
Empathy	.35	.31	.71
Motivation	..	.32	.70
Flexibility	..	..	.58

A MEASURE OF CONSEQUENCES OF ADAPTABILITY

The assessment of parent adaptability—or for that matter, any skill or ability—is of little value unless there is some means of evaluating the consequences of that ability.

One of the principal interests in the present research is the effect of parents on the growth and development of retarded children. The focus is on changes in children's social competency over a period of time. At present, we have data relative to changes over a one-year period; at the conclusion of the project we will have additional information covering a second year.

The instrument used for assessing such changes is the San Francisco Social Competency Scale.[7] This scale provides an objective check list by which a parent rates a child on a number of behaviors under four classifications: self-help, initiative-responsibility, social skills, and communication. The specific behaviors which make up the items of the scale were judged by groups of educators and psychologists, as well as parents, to be critical to the child's adequate functioning in the home situation. The reliability coefficients between mothers and fathers on independent ratings of the child were in the .80's. Reliability for mother ratings (fathers did not participate in this phase of the project) over a period of a year was .84. It seems, therefore, that this scale represents a sufficiently objective and stable measure of certain critical behaviors for retarded children.

[7] This scale was developed under the auspices of the U.S. Office of Education at San Francisco State College for the research project in mental retardation mentioned above. A discussion of its technical properties would prove lengthy and would be peripheral to the purposes of this paper. The interested reader may find this information in Leo Cain, Samuel Levine, *et al.*, "Study of the Effects of Special Day Training Classes for the Severely Retarded Child." Report to the Commissioner of Education, U.S. Department of Health, Education, and Welfare, June 1958 and June 1959.

Table 4 reports that highly adaptable mothers have children who indicate greater growth over a year's period on social competency than do middle adaptable mothers. Low adaptable mothers were eliminated from this analysis for reasons which will be discussed subsequently. Fathers were also eliminated because they did not participate in rating their children after one year.

TABLE 4

RELATION OF HIGH AND MIDDLE ADAPTABLE MOTHERS TO CHANGES IN CHILDREN'S SOCIAL COMPETENCY RATINGS OVER A ONE-YEAR PERIOD

Changes in Children's Social Competency Scores	High Adaptable Mothers		Middle Adaptable Mothers	
	N	%	*N*	%
High (+15 and above)	14	48	11	26
Middle (0 to +14)	10	35	13	30
Low (−1 and below)	5	17	19	44
Total	29	100	43	100

$\chi^2 = 6.44^*$

* $p < .05$ for 2 *df*.

The principal conclusion that can be reached from the above table is that retarded children of highly adaptable mothers show greater growth on social competency than mothers ranked as middle adaptable. It must be remembered that we have no independent way of measuring adaptability; it merely represents the sum of the component scores for flexibility, empathy, and motivation. The proposition that adaptability consists of the combination of these three factors seems to be supported by the intercomponent correlations reported in Table 3.

The fact that the criterion measure for adaptability was associated with change in the social competence of severely retarded children tends to strengthen the generalization that parent adaptability is a factor in child development. Maturation in such children is slow, and changes over a year's time are small, compared to normal children of similar chronological ages (5 to 12). It is conceivable, therefore, that, given a relation between parent adaptability and growth in such children, this relation could be as great or greater with children whose growth potential is not so severely restricted.

Previous research has emphasized the role of social class variables in differentiating parents' childrearing practices (6, 11). These results suggest the possibility that the relation between mother adaptability and change scores in children's social competency is spurious, reflecting the association of both these variables with social class. Present results, however, indicate that mothers' adaptability is not highly correlated with social class variables. As indicated in Table 5, correlations of mother adaptability and family in-

come, husband's occupational prestige, and mother education are not statistically significant. The distribution within each of these social class variables was adequate to yield a correlation if one had existed.

On the basis of these data, we can conclude that the relation between mother adaptability and change scores in children's social competency is independent of social class.

TABLE 5

MOTHER ADAPTABILITY CORRELATIONS (GAMMA) WITH SOCIAL CLASS VARIABLES

	Family Income ($N=77$)	*Husband's Occupational Prestige** ($N=78$)	*Mother's Education* ($N=77$)
Mother Adaptability . .	.22	—.07	.11

NOTE.—N's differ because of missing data.

* Derived from the breakdown of the N.O.R.C. data in Kahl (9).

In addition to the above relations, our findings shed light on some characteristics of low adaptable parents. It was found that such parents, as a group, tend to be less consistent in rating their children on social competency than do middle or highly adaptable parents. The ratings of low adaptable parents were associated with more extreme change scores in both directions over a year's period than were ratings of parents in the other categories. An F test was computed to determine the extent of the difference in variability of the social competency ratings between the lower and upper quartiles of mothers' adaptability scores. The F ratio was 2.65 in the predicted direction with 21 and 23 degrees of freedom, significant at the .05 level. As we interpret this finding, it appears that the low scoring parents, perhaps because of less involvement and less understanding of their children, were not able to rate accurately their children in terms of specific behavioral referents. The result was the greater variance and inconsistency in the ratings.

On the basis of these findings, we questioned the accuracy of the social competency ratings of low adaptable mothers. As Table 6 illustrates, a chi square computed between three classifications of children's change scores and all three groups of mothers not only reduces the significance in the difference between these groups, but also portrays the randomness of the low adaptability mother ratings. This finding in itself would be of little significance. However, when it is added to the meaningful differentiations which can be made by examining middle and high adaptable mothers' ratings, and when we consider the difference in variability in the ratings of low and high mothers, the conclusion of inaccuracy, inprecision, and in-

TABLE 6

RELATION OF MOTHERS' ADAPTABILITY TO CHANGES IN CHILDREN'S SOCIAL COMPETENCY RATINGS OVER A ONE-YEAR PERIOD

Changes in Children's Social Competency Scores	MOTHERS' ADAPTABILITY SCORES					
	HIGH		MIDDLE		LOW	
	N	%	*N*	%	*N*	%
High (+15 and above)	14	48	11	26	11	37
Middle (0 to +14)	10	35	13	30	10	33
Low (−1 and below)	5	17	19	44	8	27
Total	29	100	43	100	29	100
			$\chi^2 = 6.93^*$			

* $p < .20$ for 4 *df*.

consistency among low adaptable mothers appears to gain some support from these data.

SUMMARY

The writer in this paper has attempted to delineate some of the theoretical and methodological considerations underlying the development of an instrument designed to assess parents' problem solving abilities in child-rearing situations. Such skills are conceived of as including, in combination, the attributes of empathy, flexibility, and motivation. This complex of components is considered to determine the individual's ability to deal effectively with problematic situations by changing roles, behaviors, and attitudes when required, a skill which has been labeled here as adaptability.

Results of the research reported here have shown a significant relation between parents' adaptability and growth in the social competency of retarded children. In addition, evidence is presented to support the contention that low adaptable parents are less precise, consistent, and accurate in their perception of their children's behavior.

Mother-father differences in adaptability have not been discussed in this paper, although there is evidence indicating that such differences exist. Analysis and additional information is being compiled on this subject. A discussion of adaptability as related to parental roles has been deleted principally because the possible explanations and generalizations are inclined to be limited to questions of rearing a retarded child, and therefore, not within the scope of this paper. The possibility exists, however, that questions of the nature of a parent's identification with his child may tend to specify the conditions under which a parent may be more or less adaptable.

The parent adaptability instrument attempts to assess childrearing within the parent's personal experiential and normative framework. This is done

by considering the differential use of cognition and specificity with which the parent confronts and responds to a situation, rather than examining specific behaviors. This approach offers the possibility for re-evaluating the implications of social class differences in childrearing (6, 11). For example, the differences reported by Maccoby and Gibbs (11) for middle and working class parents in such areas as techniques of discipline may be reinterpreted from a problem solving point of view. The consequences of using different techniques of discipline may vary in relation to the expectations and meaningfulness of the act to the child and the parent. It may be as ineffective for a middle class parent to be physically punitive as for a working class parent to use withdrawal of love as a disciplinary measure. If we ignore the specific type of act employed, but, rather, approach the situation in terms of the appropriateness of the act based on the parent's understanding of the child and the situation and the parent's involvement in doing something about the situation, our inference may be different. As a matter of fact, present results indicate that, for mothers, adaptability is not highly correlated with social class variables.

The procedure for determining adaptability depends upon a classificatory system of measurement, which is lacking in precision and involves a considerable amount of error. Nevertheless, the present research indicates that it is capable of producing fruitful results in assessing childrearing practices. In addition, the concept of adaptability may have heuristic implications in assessing the parental contribution to behavioral disorders in children.

References

1. Allport, G. W. The historical background of modern social psychology. In G. Lindzey (Ed.), *Handbook of social psychology*. Addison-Wesley, 1954. Pp. 3-56.
2. Baldwin, A. L. Role of an ability construct. In D. C. McClelland (Ed.), *Talent and society*. Van Nostrand, 1958. Pp. 195-232.
3. Baldwin, A. L., Kalhorn, J., & Breese, F. H. Patterns of parent behavior. *Psychol. Monogr.*, 1945, 58 (Whole No. 268).
4. Bruner, J. S., & Tagiuri, R. The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology*. Addison-Wesley, 1954. Pp. 634-654.
5. Burgess, E. W., & Wallin, P. *Engagement and marriage*. Lippincott, 1953.
6. Davis, A., & Havighurst, R. J. Social class and color differences in child rearing. *Amer. sociol. Rev.*, 1946, 11, 698-710.
7. Dymond, R. A scale for the measurement of empathic ability. *J. consult. Psychol.*, 1949, 13, 127-133.

8. Frenkel-Brunswick, E. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, 18, 108-143.

9. Kahl, J. A. *The American class structure.* Rinehart, 1957.

10. Lindzey, G., & Borgatta, E. F. Sociometric measurement. In G. Lindzey (Ed.), *Handbook of social psychology.* Addison-Wesley, 1954. Pp. 405-448.

11. Maccoby, E. E., & Gibbs, P. K. Methods of child-rearing in two social classes. In W. E. Martin & C. B. Stendler (Eds.), *Readings in child development.* Harcourt, Brace, 1954. Pp. 380-396.

12. Murphy, G. *Social motivation.* In G. Lindzey (Ed.), *Handbook of social psychology.* Addison-Wesley, 1954. Pp. 601-633.

13. Newcomb, T. *Social psychology.* Dryden, 1950.

14. Notcutt, B., & Silva, A. L. Knowledge of other people. *J. abnorm. soc. Psychol.*, 1951, 46, 30-37.

15. Parsons, T., & Bales, R. F. *Family, socialization and interaction process.* Free Press, 1955.

16. Rokeach, M. *The open and closed mind.* Basic Books, 1960.

17. Rommetveit, R. *Social norms and roles.* Univer. of Minnesota Press, 1958.

18. Sarbin, T. R. Role theory. In G. Lindzey (Ed.), *Handbook of social psychology.* Addison-Wesley, 1954. Pp. 223-258.

19. Schaefer, E. S., & Bell, R. Q. Development of a parental attitude research instrument. *Child Develpm.*, 1958, 29, 339-361.

20. Sears, R., Maccoby, E. E., & Levin, H. *Patterns of child rearing.* Row, Peterson, 1957.

21. Strodtbeck, F. L. Family interactions, values, achievement. In D. C. McClelland (Ed.), *Talent and society.* Van Nostrand, 1958. Pp. 135-194.

22. Thomas, W. I., & Thomas, D. *Child in America.* Knopf, 1928.

23. Torgerson, W. *Theory and methods of scaling.* Wiley, 1958.

24. Whiting, J. W. M., & Child, I. L. *Child training and personality: a cross-cultural study.* Yale Univer. Press, 1953.

CHILDREARING ATTITUDES OF MOTHERS IN GERMANY AND THE UNITED STATES[1]

DON W. RAPP*
University of Illinois

Information concerning the relationship between childrearing attitudes in Germany and the United States has been, until recently, obtained largely from nonempirical data. Lewin (8), as an expert observer, expressed the view that German children are expected to exhibit unquestioning obedience, while children in the U. S. are urged to be more independent. Klineberg (6), however, has remained cautious in committing himself to statements regarding the German family's insistence on obedience. He used the phrase "allegedly characteristic" in reference to the reported widespread emphasis on obedience in the German culture. Supporting Lewin's views with direct cross-national investigations, Anderson (1), McGranahan (9), and Schaffner (11) used German and U. S. youth and young adults in detecting national differences with respect to a more highly developed authority structure in the German culture.

Cohn and Carash (2), using the California F Test, found a German sample of working men more "authoritarian" than a similar U. S. sample. The German sample had a higher authoritarian score on the F Test than any national sample hitherto reported. Kates and Diab (5) obtained a positive, although nonsignificant correlation, between the F Test and childrearing attitudes; i.e., the more authoritarian the more controlling the attitudes concerning children.

The present study sought to investigate directly whether the childrearing attitudes of German and U. S. mothers were empirically distinguishable.

METHOD

Subjects

German subjects were mothers from the Länder of Baden in southwest Germany, who responded to a German translation of the University of Southern California Parent Attitude Survey (12), PAS, and an attached personal information sheet. The sample was obtained through a Protestant

* Division of Child Development, University of Illinois, Urbana.

[1] A condensation of a dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Florida State Univer., August, 1960. The author wishes to express his gratitude to his graduate committee: Drs. Ruth J. Dales, Ruth Connor, Sarah Lou Hammond, Ralph L. Witherspoon, and James Walters.

Child Develpm., 1961, 32, 669-678.

kindergarten organization.[2] The investigator met with the 30 kindergarten teachers who served as liaison persons for the study. Distribution of 200 PASs was begun in May 1958, and in the same year 183 or 91.5 per cent were returned to the investigator in the U. S.

The PASs in the U. S. were distributed in 1960 to 28 Florida preschools in the same general manner as in the German sample. Of the 537 PASs distributed, 283 or 52.7 per cent were returned. The difference in the percentage of survey return between the German and U. S. samples was thought in part to be a function of the unifying organization to which the entire group of German teachers belonged. The U. S. teachers involved in this study lacked any common organizational ties.

The "kindergarten" in Germany includes a wider age range than the U. S. kindergarten; therefore the U. S. sample was drawn from mothers of kindergarten and nursery school children.

All subjects in both final samples were Caucasian, Protestant, and from intact homes. They reported parents, husbands, and themselves to be citizens of their respective countries, had at least one child in preschool, and completed PASs omitting no more than three items. Weights for omitted items were interpolated from the mean of the subscale to which they belonged.

The Index of Social Status (10), ISS, used in this study was originally based on U. S. research. Lefkowitz (7), testing its appropriateness for use with a European sample, found a correlation of $r = .82$ ($p < .01$) between expert local (Vienna, Austria) social class judgment and ISS categorical placement. The ISS determines the social status of a family by rating the occupation, source of income, and educational attainment of the "status parent." In this study information concerning the father was used. The ISS divides the social class continuum into five categories, two lower and two middle levels and one upper level. Subjects in the present study who fell into the ISS upper ($N = 4$) and upper-middle ($N = 41$) class categories were combined and designated as the "upper" class subsample with an N of 45. The subjects who fell into the ISS lower middle class category were designated as the "middle" class subsample of 41 subjects, and the subjects who fell into the upper-lower ($N = 30$) and lower-lower ($N = 8$) ISS class categories were combined into a "lower" class subsample of 38 subjects. The terms upper, middle, and lower as used in this study are relative terms and should be thought of in reference to their make-up. Social class breakdown, in ISS categories, was identical in both national samples. The upper, middle, and lower class subsamples, when considered separately, are designated as class samples; combined they are called complete samples.

The German and U. S. samples were equated by individually matching the U. S. surveys with the German surveys. Of the 124 German-U.S. pairs,

[2] Sincere appreciation to Pfarrer Herrenbrodt, head of the Gesamtverband der Inneren Mission, Karlsruhe, Germany, to Fräulein Gegenheimer, Mission supervisor of Baden, and to Fräulein Maria Fath, administrative secretary of the Mission, whose combined efforts made possible the collection of the German data.

86 were matched exactly for social class, age of mother in five-year intervals, and number of children in the family. The remaining 38 pairs were matched for social class, and as closely as possible on the other two categories, neither of which was found significantly different. The mean age of the German sample was 32.96, SD 5.58; for the U. S. sample, 32.80, SD 5.78. The mean number of children per mother for the German sample was 2.3, SD 1.23; for the U. S. sample, 2.5, SD 1.33.

The education of the mothers was not controlled in the equating of the two samples. In both German and U. S. samples 48 mothers reported two or more years of college; 59 U. S. and 19 German mothers reported themselves to be at least high school graduates or the equivalent; and 17 U. S. and 43 German mothers completed 11 years of school or less. Fourteen of the German mothers failed to report their educational attainment. The χ^2 of 31.1 summarizes the educational differences between the two samples.

Instrument

Shoben's Parent Attitude Survey was utilized in German and English forms. In Shoben's original research (12), the 85-item survey discriminated between mothers of problem children and mothers of nonproblem children. A low score has been interpreted as indicating a more optimum attitude concerning child guidance; a high score indicates the reverse. Three main subscales are present in the PAS: Dominant, Possessive, and Ignoring. The Dogmatism score (3) was derived from the actual number of extreme responses, i.e., strongly agree and strongly disagree.

The PAS was translated by four bilingual Germans.[3] The investigator was present with three of the four at the time of translation. Questions were answered concerning the particular American colloquialisms which frequent the PAS. These four translations were then used by a fifth translator[4] to bring together the final German form.

Results

The complete samples were cross-nationally compared on the five scores obtained: PAS total score, Dominant, Possessive, and Ignoring subscale scores, and the Dogmatism score. All five tests were found significant beyond the .001 level with the U. S. sample demonstrating the smaller scores or the less controlling attitudes. Figure 1 presents graphically the wide differences found between the total score distributions of the two complete samples. The median test (13) χ^2 value was 80, significant well beyond the .001 level. This large difference is reflected in the statement: only 4 per cent of the U. S. mothers had total scores above the median of the complete

[3] Dr. G. Hassfurther, Dr. R. Defieber, Dr. H. Reinstrom, and Dr. M. Suessman.

[4] Dr. Marlies Kronegger.

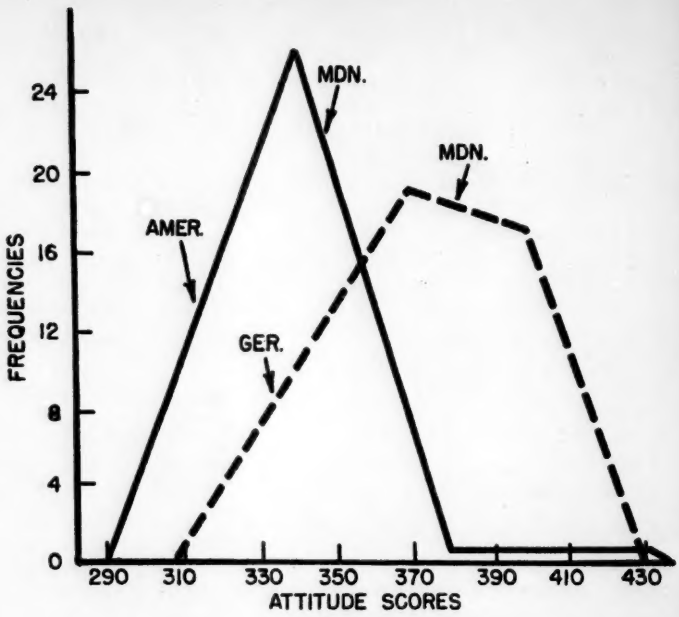

FIGURE 1—Distributions of total PAS scores from the complete German and United States samples. The median test χ^2 value was 80 ($p < .001$).

German sample. Other values on the complete sample cross-national tests were: Dominant subscale ($\chi^2 = 62.02$), Possessive ($\chi^2 = 76.94$), Ignoring ($\chi^2 = 12.75$), and Dogmatism ($\chi^2 = 19.76$). Considering the great disparity between the German and U. S. samples on all five scores tested, it can be said that the two samples were most probably drawn from attitudinal populations with different medians.

German and U. S. comparisons of like social classes—i.e., German upper-class to U. S. upper-class, etc.—were made using the Mann-Whitney U test (13). Corrections were made for ties. Figure 2 presents these comparisons in horizontal relationship on all five scores tested. All of the three U. S. social classes were found to possess smaller PAS median scores, indicative of less controlling childrearing attitudes, than their German social class counterparts. Thirteen or 87 per cent of the 15 comparisons reached significance, one at $p < .05$ and 12 at $p < .01$. The two tests not reaching significance were the cross-national upper class comparison in the Ignoring subscale and the middle class Dogmatism comparison. They were, however, in the expected direction ($p < .10$, $p < .17$, respectively). The findings

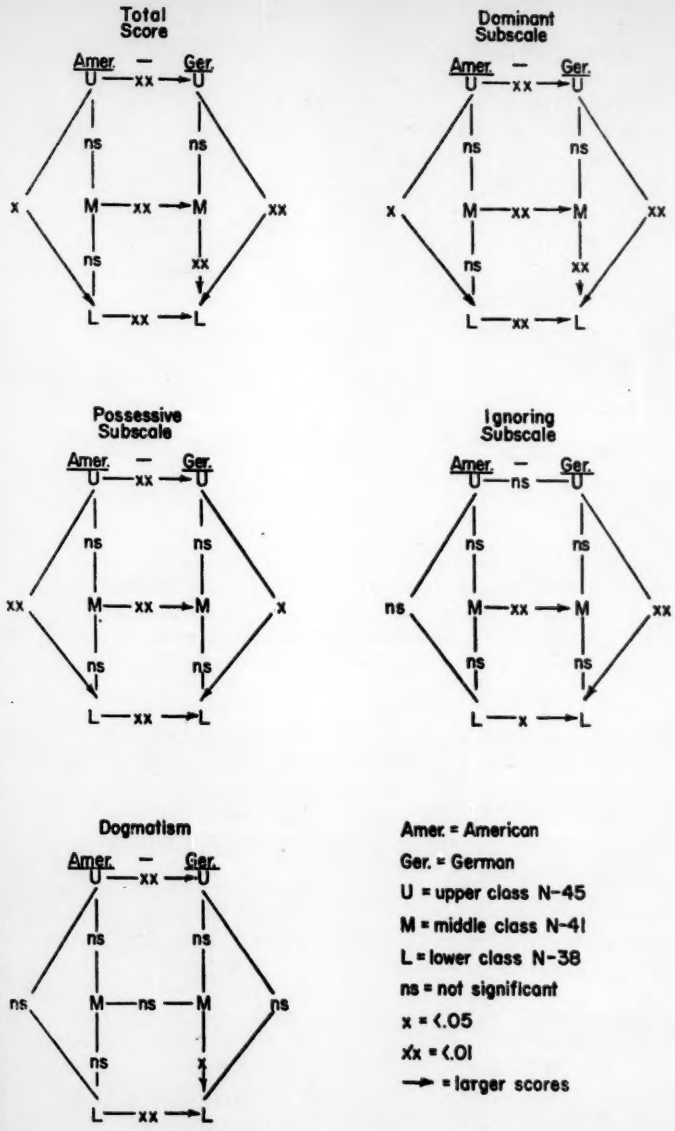

FIGURE 2—Inter- and intranational social class relationships of German and United States childrearing attitudes.

leave only a minimum of doubt that comparable German and U. S. class samples were independent samples drawn from separate populations.

Intranational class comparisons were computed between upper and middle, middle and lower, and upper and lower class samples on all five scores obtained. Figure 2 presents the intranational comparisons in vertical relationship. Twenty-seven of the 30 class comparisons indicated the following pattern: upper class scores smaller than middle and lower class scores, and middle class scores smaller than lower class scores. Ten or 33 per cent of the 30 intranational tests were found significant, four at $p < .05$ and six at $p < .01$. Significance was generally reached between upper and lower class samples in both national samples and between the German middle and lower class samples. Upper and middle class samples in both national samples responded quite similarly. Thus, it is evident that the U. S. social classes exhibited less variability than did the German social classes. The reason is primarily due to the significant differences reached between the German lower and middle classes as opposed to no U. S. differences in this relationship.

Discussion and Conclusions

A major limitation in all research crossing a language barrier is the insecurity of identical bilingual understanding of the material presented. A second limitation should also be noted. The German situation made it impossible to pretest the German translation. The manner in which the survey was translated, however, was designed to overcome these limitations to the greatest degree possible.

The PAS in the original study discriminated between mothers of problem children and mothers of nonproblem children. German scores approximated the scores obtained by mothers of problem children, and the U. S. scores approximated the scores obtained by mothers of nonproblem children. This should not be interpreted as indicating that German children have a greater chance of becoming problem children than U. S. children. In this respect the uniqueness of culture does not allow judgmental interpretation.

The German mothers reported significantly less education than the U. S. mothers. The matching inaccuracy stems from the fact that the ISS determines social class by using criteria based on the husband. The relationship indicated here is an interesting one. The educational level of the U. S. mothers in this sample was similar to that of their husbands, while the educational level of the German mothers was much less than their husbands. Relevant to the present findings, it has been found (12, 15, 17) that the more education a person has the less controlling their childrearing attitudes. It would then seem possible that part of the large score difference found in the present study could possibly be attributed to the difference in educational level of the two national samples. Analysis of this possibility was conducted as follows: The total scores of the U. S. lower class (least edu-

cation of the U. S. sample) were compared with total scores of the German upper class (most education in the German sample). Differences were still found,[5] indicating the U. S. lower class had significantly smaller scores or less controlling attitudes than the German upper class.

The three U. S. social classes exhibited less variability than did the German social classes (Figure 2). These data are interpreted as a reflection of cultural influences. Specifically, greater social class uniformity in the U. S. sample is seen, at least in part, as the result of the American governmental system advocating equality and the public educational system which promotes contact between social classes. Greater German social class disunity is seen partly as a result of the very recent dictatorial militaristic form of government and a more authoritarian class determining educational system. This finding is in support of McGranahan's (9) statements indicating that German-U.S. differences in attitudes are not peculiar to political attitudes, but spread throughout a wide variety of attitudinal areas into even the simplest types of social relationships.

Comparison of PAS Scores

Table 1 has assembled the means and SDs of several studies which have used the complete form of the PAS.[6] The value of placing present study results in the context of previous investigations is beneficial not only to the present research but allows multiple comparison. The German total and Dominant mean scores are considerably higher than all the scores presented except for Shoben's so-called problem group. As shown in this study, the U. S. sample demonstrated significantly smaller scores than the German sample. However, the U. S. sample also has total and Dominant means that are considerably higher than all but those for Shoben's problem group.

Specifically, all three of the PAS means presented by Walters *et al.* (16) were obtained from samples of mothers who generally belonged to the upper middle class. It was thought possible that this class restriction in the "three state sample" (16) may have been one of the reasons why the present study U. S. sample, with its full range of social class, tended to exhibit higher scores. However, when the mean of the upper class U. S. sample (344.2, SD 15.9, N 45) was compared with the mean of the Florida, Alabama, and Connecticut means (328.1, approximated SD 26,[7] N 90), a significant difference was found well beyond the .01 level. Direction and

[5] Mann-Whitney U test $z = 3.67$; $p < .001$. Two-tailed tests of significance were used throughout the analysis.

[6] Only 84 of the 85 PAS items were used in this study. Item 1 of the Ignoring subscale was omitted from the scoring of both samples. A typographical error on the German form rendered it dissimilar in meaning to the English version. In order to make these data of the present study comparable to the other PAS data presented, four score points (an approximation of the mean weight responses on the omitted item) have been added to the total and Ignoring scores in both samples.

[7] The approximated SD of 26 is presumed a reasonable if not a high estimate of the true SD considering the homogeneous social class nature of the "three state sample" and also considering the range and magnitude of the other SDs presented in Table 1.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF PAS SCORES OBTAINED IN COMPARABLE STUDIES

	Date of Data Collection	Total		Dominant		Possessive		Ignoring	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Present Study									
American mothers	1960	†352.1	26.0	169.3	13.1	82.8	8.4	† 56.5	5.2
German mothers	1958	385.3	24.9	186.6	13.9	94.7	7.7	† 59.5	5.1
Shoben									
Nonproblem mothers (Gp I)	1949*	324.7	21.7	154.7	11.6	91.1	6.4	42.8	4.2
Problem mothers	1949*	397.7	29.4	187.8	16.7	111.7	10.3	56.1	5.9
Clinical psychologists	1949*	286.4	6.7	146.1	2.2	79.4	4.6	39.9	2.7
Freeman									
Nonschiz. mothers	1955*	. . .	. . .	158.5	18.1	85.9	10.7	56.8	5.4
Schiz. mothers	1955*	. . .	. . .	161.0	16.7	89.6	11.3	58.6	6.7
Walters et al.									
Florida mothers	1957	322.2	. . .	154.6	. . .	78.0	. . .	54.4	. .
Alabama mothers	1955	333.7	. . .	157.3	. . .	77.3	. . .	55.6	. .
Connecticut mothers	1957	328.6	. . .	156.7	. . .	77.2	. . .	54.0	. .
Walters & Fisher									
September, 1953	1953	338.3	. . .	158.0	. . .	82.8	. . .	54.9	. .
January, 1954	1954	326.9	. . .	152.0	. . .	78.9	. . .	59.6	. .
November, 1955	1955	311.4	. . .	146.0	. . .	76.4	. . .	49.7	. .
Walters & Bridges									
Men: Average of seven college samples	1955	337.7	. . .	159.6	. . .	82.9	. . .	52.2	. .
Women: Average of two college samples	1957*	338.6	. . .	160.0	. . .	82.6	. . .	54.3	. .

* Data collected prior to date indicated.

† Four score points have been added in order to render the means comparable to the other means presented (*see* footnote 6).

magnitude of this comparison tends to lower the probability of social class completely explaining the mean score difference.

U. S. data of the present study were collected in 1960 while data from other American studies were collected in 1957 or before. The present study U. S. total and Dominant subscale scores appear higher than any of the other mean scores here presented, with the exception of that for Shoben's problem mother sample. The question arises, "Within the last few years are the mothers in the United States becoming more dominant and controlling over their children?" If this is evidence of a general shift to a more controlling childrearing attitude, it is one that coincides with the coming of the sputnik era with its re-emphasis on the three R's and accompanying public re-evaluation of educational techniques and curricula.

Summary

The purpose of this investigation was to determine the measurability and relationship between German and U. S. childrearing attitudes. Parent Attitude Surveys from 124 Florida mothers were individually matched with translated surveys completed by 124 German mothers from Baden in southwest Germany. Matching was near perfect on three categories, social class, age, and number of children in the family.

Significant differences were found between the German and U. S. complete samples on all five scores obtained. The U. S. mothers demonstrated scores indicating less controlling attitudes than the German mothers. These differences were maintained in 13 of the 15 cross-national comparable class comparisons. The analysis within social class showed similar relationships in both samples. Upper and middle class scores were very much the same within each national sample; however, the lower classes demonstrated the most controlling and authoritarian attitudes within their respective cultures. Social classes in the U. S. were found attitudinally less variable than those in Germany with respect to childrearing attitudes.

A comparison of available Parent Attitude Survey scores from other studies was presented.

References

1. Anderson, H. H., Anderson, G. L., Cohen, I. H., & Nutt, F. D. Image of the teacher by adolescent children in four countries: Germany, England, Mexico, United States. *J. soc. Psychol.*, 1959, 50, 47-55.
2. Cohen, T. S., & Carsch, H. Administration of the F scale to a sample of Germans. *J. abnorm. soc. Psychol.*, 1954, 49, 471.
3. Freeman, R. W., & Grayson, H. M. Maternal attitudes in schizophrenia. *J. abnorm. soc. Psychol.*, 1955, 50, 45-52.

4. JAHODA, M., DEUTSCH, M., & COOK, S. W. *Research methods in social relations.* Dryden Press, 1958.

5. KATES, S. L., & DIAB, L. M. Authoritarian ideology and attitudes on parent-child relationships. *J. abnorm. soc. Psychol.*, 1955, 51, 13-16.

6. KLINEBERG, O. Cultural factors in personality adjustment of children. *Amer. J. Orthopsychiat.*, 1953, 23, 465-471.

7. LEFKOWITZ, M. Authoritarianism in adolescents: a cross-national study. Unpublished doctoral dissertation, Univer. of Texas, 1956.

8. LEWIN, K. *Resolving social conflicts.* Harper, 1948.

9. MCGRANAHAN, D. V. A comparison of social attitudes among American and German youth. *J. abnorm. soc. Psychol.*, 1946, 41, 245-257.

10. MCGUIRE, C., & WHITE, G. D. The measurement of social status. Research paper in human development, No. 3 (Rev.). Univer. of Texas, Austin, 1955.

11. SCHAFFNER, B. *Fatherland: a study of authoritarianism in the German family.* Columbia Univer. Press, 1948.

12. SHOBEN, E. J., Jr. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.

13. SIEGEL, S. *Non-parametric statistics for the behavioral sciences.* McGraw-Hill, 1956.

14. WALTERS, J., & BRIDGES, B. Attitudes of single men toward child guidance. *J. home Econ.*, 1956, 48, 109-113.

15. WALTERS, J., & FISHER, C. Changes in the attitudes of young women toward child guidance over a two-year period. *J. educ. Res.*, 1958, 52, 115-118.

16. WALTERS, J., STRICKLAND, M. P., & CHRISTOPHERSON, V. A. Paternal and maternal attitudes concerning the guidance of children. Unpublished manuscript, Dept. Home & Family Life, Florida State Univer., 1959.

17. ZUCKERMAN, M., BARRETT, B. H., & BRAGIEL, R. M. The parental attitudes of parents of child guidance cases. *Child Develpm.*, 1960, 31, 401-417.

NEW TECHNIQUES FOR RESEARCH IN CHILDREN'S TEMPTATION BEHAVIOR

ROBERT E. GRINDER*
University of Wisconsin

The observation that progress in a scientific field usually waits upon the discovery of a satisfactory dependent variable (12) seems particularly relevant in regard to temptation behavior. The term "temptation" is used here to refer to a type of conflict where a choice must be made between (a) conformity to some socially expected role behavior and nonattainment of a goal and (b) nonconformity and attainment of the goal.

For several years, on interview, case history material, and theoretical bases, personality theorists have viewed the kind of response(s) that an individual makes to temptation as one of the indices of the construct conscience or superego. For example, Sears *et al.* (11) suggest that capacity to resist temptation is one of the consequences of conscience development. Freud speaks of the superego as generating moral anxiety as it "observes, guides, and threatens the ego" (6, p. 89). Fenichel describes the "warning function" of conscience and distinguishes guilt feelings from "feelings which do not judge the past but the future" (4, p. 134). Redl and Wineman, in discussing "value danger signals," suggest that the superego uses the production of feelings of guilt as either "pre-situational or post-situational signals of value conflict to the ego" (10, p. 200). The contribution of these and similar studies to theoretical knowledge and empirical ways of assessing the conscience or superego construct is valuable; therefore, it is necessary that operational definitions of responses to temptation be developed in order to test them as dependent variables in empirical investigations of superego functions. Because research method in this area is by and large uncharted, this paper attempts to outline some of the general theoretical problems and to provide a methodology which has been demonstrated as a measure of children's temptation behavior (3, 7).

THEORETICAL FRAMEWORK

The common-sense notion of temptation is offered by the Oxford Dictionary (13) which presumes a person to be tempted when he is "enticed to do what is wrong by the prospect of some pleasure or advantage." For empirical research, however, it seems essential that the meaning of the word be specified in a strictly functional or behavioral sense (cf. Bijou's [2]

* School of Education, University of Wisconsin, Madison 6.

Child Develpm., 1961, 32, 679-688.

approach to an empirical behavior theory for children). A definition of temptation of this kind may be conceptualized as follows: a state of temptation exists for an individual when the addition or presence of a rewarding stimulus (incentive) increases the probability of a class of responses which is forbidden because of its incompatibility (conflict) with some socially expected behavior.[1] This definition is based upon the functional relationships between the elements of a temptation situation itself.[2] The incentive for temptation conceivably might be an intangible such as the anticipation of praise or the circumventing of anticipated reproof, or a tangible such as money, candy, a toy, or freedom from corporal punishment. The behavior upon which transgression is contingent might be evaluated from the point of view of virtues, e.g., honesty, truthfulness, punctuality, responsibility, etc., or from that of tabus, e.g., cheating, lying, procrastination, carelessness, etc. According to the definition, then, a tempted individual is in the sort of conflict that Miller (9) has named "double approach-avoidance"; both alternatives, either yield to temptation or resist temptation, appear as approach tendencies, but the necessity of relinquishing one of the alternatives in order to pursue the other may instigate a tendency to avoid the one which is approached.

Whatever the character of the conflict, the probability of transgression, and thus the degree of conflict, (a) will vary inversely with extent of coercion and/or the possibility of being detected and will vary directly with (b) the extent to which the expected behavior has been learned and (c) the incentive value of the tempting objective (cf. below). Further, (d) the objective of temptation may be either an acceptable goal which can be obtained only by means which are forbidden or a goal which is forbidden. Finally, (e) temptation conflict, in the sense that the person is pushed in two or more directions simultaneously (5), ceases when the desired goal has been achieved or when it has been successfully resisted.

From the foregoing presentation it may be concluded that temptation is a predecision-type conflict situation where an individual, free from coercion or danger of detection, must choose between a positive incentive and conformity to learned role behavior.

If subjects are to be compared with one another in the degree to which they either resist (conform) or yield (transgress) to temptation, it is necessary that the variables of the temptation conflict be measurable indices. The

---

[1] Socially expected behavior is conceptualized to be a specific aspect of role behavior, e.g., a child assigned to be class monitor might transgress this role by throwing a spitball, deserting his post, or napping, etc. In an experiment socially expected behavior may be defined by consensus of qualified judges.

[2] It is recognized nonetheless, that in seeking developmental antecedents of temptation behavior, as for example in aspects of socialization (1, 3, 7), it might be useful to postulate personality constructs (e.g., conscience or superego) or to infer such motivational properties as drives, desires, wishes, etc. Further, on this level of analysis, it might be appropriate to introduce such intervening constructs as identification, introjection, habit, etc.

expected behavior pattern provides a single continuum upon which the variables can be ordered. If a subject conforms to the pattern, i.e., if he behaves as he would under nontemptation conditions, he can be scored as having resisted temptation. There is, in other words, zero discrepancy between his behavior and the expected behavior. On the other hand, if he transgresses, his responses of deviation can be taken as measures of yielding to temptation. There is, in this instance, a discrepancy between the subject's behavior and the expected behavior.

These indices also are not independent of each other. The conflict is a function of the attraction of the tempting goal *and* the extent to which the expected behavior has been learned. Consider the problem, important to studies of self-control, of comparing subjects in terms of resistance to temptation: some subjects may appear to resist temptation whereas actually the incentive is not tempting, and others may apparently transgress, not because they have not learned the expected behavior as well as those who have conformed, but because the incentive has relatively more appeal to them (7). Obviously, only when the tempting object has equal incentive value to all subjects can the extent to which they conformed be meaningfully compared.

Controlling the incentive value of the tempting goal is difficult. The tempting objective, although on occasion intangible, is usually some tangible, manipulable goal, the incentive value of which must be inferred from goal-directed behavior in nontemptation, nonconflict situations. Also, the relative desirability or reward properties of goal objects, e.g., candy, toys, or money, etc., must be demonstrated for each individual. The great difficulty of these tasks is illustrated by the fact that in the now classical Hartshorne and May (8) Character Education Inquiry efforts to control temptation were abandoned. In the methodology presented here, a simple procedure for controlling the incentive value of the tempting goal proved itself to be useful. The same incentive stimuli were offered to all subjects in an experiment, and an attempt was made to make the incentive attractive so that every subject's motivation would be high.

In addition to settling upon suitable indices of the dependent variables and assessing temptation to transgress, in devising an experimental task a number of technical considerations are necessary. First, so that activity in the situation will be appealing, interesting, provocative, meaningful, etc., experimental circumstances should be appropriate to the subjects' sophistication. That is, the activity should be sufficiently complex as to arouse no suspicion or fear of external coercion and/or detection, while at the same time transgression is relatively easy. Second, achievement or progress toward the goal should be constant in order to assure that the tempting goal is equally attainable by all subjects. Otherwise, some subjects, skilled in the task, may not be tempted to transgress because the goal can be attained readily. Third, since temptation to transgress may be contingent upon knowledge of the goal's presumed attainability, subjects should have opportunity to keep a cumulative record of their progress toward the goal.

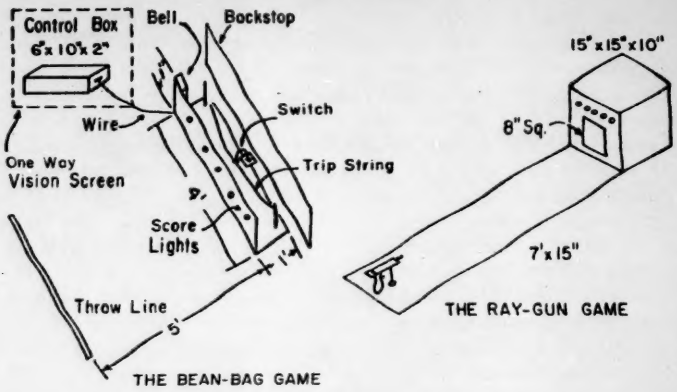

FIGURE 1—The apparatus for measurement of temptation behavior.

Fourth, an objective recording of the subjects' behavior should be made. Fifth, the apparatus should be of a sort that is easy to transport and operate, because of practical difficulties in bringing subjects to a laboratory and because fear of detection is minimized in a familiar setting.

With these considerations in mind, two temptation situations have been devised, a beanbag game for children 4 to 5 years old and a ray-gun game for those from 6 to 15 years old (Figure 1).[3] A game situation was chosen for both age groups for several reasons. First, it seemed likely that each child would have learned in varying degrees that in a game his behavior should be to "play fair" or to "play according to the rules." Second, the expected behavior can be specified in rules easily understood by children. Finally, it is likely that to some extent this aspect of socialization would have been introduced relatively early in the child's training. The measurement procedure for both age groups generally is the same. Subjects, asked to play individually, are tempted to transgress its rules in order to win an appealing incentive.

METHODOLOGY

The Beanbag Game

Procedure. A subject is individually instructed by an experimenter to stand behind, and not to cross, a throw-line drawn parallel 5 ft. from a 1-

[3] The methodology presented here and the essential ideas for the construction of the beanbag game grew out of discussions at a graduate seminar at the Laboratory of Human Development at Harvard. While the staff of the Laboratory and other members of the seminar should not be held responsible for this particular formulation, the basic ideas are a result of our joint effort. The writer wishes to acknowledge in particular the guidance received from John W. M. Whiting, Wesley Allinsmith, Eleanor E. Maccoby, and Roger V. Burton.

by 4-ft. panel upon which five 1-in. pilot (score) lights are evenly spaced across the front. The subject is given five beanbags which he is to toss one at a time over the panel. If a bag "scores," it turns on a light. None of the bags is to be retrieved. The subject is told that if he can turn on three lights he will win an exciting toy; he is then shown a bowl full of 25 cent toys from which he is asked to choose the one he would like most to have if he should win.

To insure that the subject understands the rules, he is asked to play a practice game. As the next game, the one for the toy, is about to begin, the experimenter announces that he will have to attend to a telephone call. The subject is then asked (because the experimenter will have to be away for a few minutes) to play the game by himself. The scores each subject earns by tossing the bags are controlled by an observer behind a one-way vision screen. If the subject follows the rules, only one light will register. However, if he breaks the rules, a light will register on each occasion that an unauthorized bag is tossed. The subject is judged to have resisted temptation if he conforms to the rules; he is judged to have yielded to temptation if he transgresses. The latter may be measured in terms of how many bags are tossed before the subject transgresses or in terms of the number of extra tosses or the number of times the throw line is crossed, etc.

After sufficient time has elapsed for the child to play the game, the experimenter returns to the game room. At this point, at least two alternatives are possible. He may express regret that the child who conforms did not win a toy, and he may reward the child who has transgressed with the toy of his choice. On the other hand, as Burton *et al.* (3) chose to do, he might suggest that the game be played again for the toy. Now the observer turns on four lights as the bags are tossed. Every child departs with a toy, and an opportunity to relieve possible guilt is offered to those who have transgressed.

Technical description. Figure 2 shows the schematic circuit of the beanbag game. The dashed line separates the control box from the game. A multiconductor cable, which runs along the floor to behind the one-way vision screen, connects the two units. Notice that the control box houses a circuit on-off switch (S_1), a six position rotary-shorting switch (S_2) for the score lights, a push-button switch (S_3) for ringing a chime every time a score light registers, a score circuit switch (S_4) parallel to a similar switch (S_5) on the front panel, and a score indicator lamp. The game, built of plywood, is approximately 4 ft. long with the panel and its base about 12 in. in height and width, respectively. The five score lights are evenly spaced along the bottom of the front panel. A switch (S_5), mounted in the upper left-hand corner of the front panel, enables the experimenter to turn off the score lights after an experimental sequence. When S_5 opens the circuit, the observer at the control box can return the rotary switch "home," and then close the circuit again with S_4. A string, which passes through a loop on the end of a stem of a microswitch (S_6), serves as a target for the

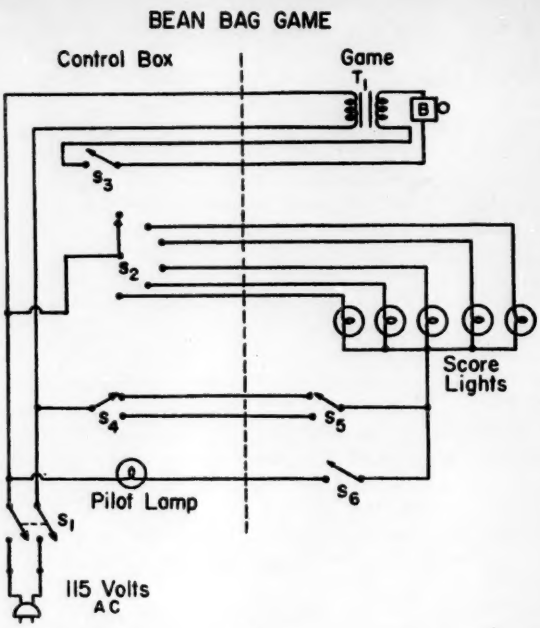

FIGURE 2—Schematic circuit of the beanbag game. S_1 = D.p.s.t. toggle. S_2 = 6-pole 1-position rotary shorting switch. S_3 = S.p.s.t. push button. S_4, S_5 = S.p.d.t. toggle. S_6 = S.p. momentary position microswitch. T_1 = Bell transformer, 115 v., 6.3 v. B = Doorbell, 6.3 v. Bulbs = 6 w., candelabra base, 115 v.

beanbag tosses. Whenever the switch is momentarily closed by a bag hitting the string, a score indicator lamp on the control box flashes. The string, mounted behind the front panel and out of sight of the subject, runs the length of the game; S_6 is mounted in the center of the base. Several modifications of the apparatus are obviously possible. Burton *et al.* (3), for example, substituted SPST mercury switches wired in series for the control box rotary switch in order to eliminate its contact noises; also, they built a portable one-way vision screen which made it possible to use the game in any available school room.

The Ray-Gun Game

Procedure. The game consists of a "ray-gun," a 7 ft. by 15 in. plank (hinged together in three sections) upon which the gun is swivel mounted,

and a 15 by 15 by 10 in. plywood box. On the front of the box are two revolving illuminated rockets set behind a piece of translucent glass, 8 in. square and five 1-in. pilot lights evenly spaced across the front of it above the target glass. Scores from 0 to 5 are registered on the pilot lights with each pull of the trigger of the ray-gun. Obtained scores are prearranged (cf. technical description below).

The experimenter demonstrates, for about 10 subjects at a time, the game and the rules for playing it. Subjects are told that if they "hit" the center of one of the rockets they will score a bull's-eye and all five of the score lights will shine; if they hit the head or the tail of the rocket they might get a score from 0 to 5 depending on how close they are to the center. Subjects are asked to take 20 shots and to record their scores on a three-column, 20-row score sheet with a ball point pen. The first column, labeled "shot number," lists the row number; the second column, labeled "score obtained," is blank in order that a subject can record whatever score he might choose for each shot; the third column, labeled "total so far," is for a subject to record cumulatively his scores.

Subjects are told that if they obtain from 35 to 39 points they will win a Marksman badge, 40 to 44 points a Sharpshooter badge, and 45 points or more an Expert badge. As has been noted, scores are prearranged, and the experimental sequence totals 32 points when *only* obtained scores are recorded. The badges, 1¾ in. in diameter, were prepared by a commercial badge firm which engraved in black a drawing of the ray-gun upon a background colored differently for each of the three types of badge.[4]

After explaining the rules, the experimenter "aims" carefully, takes a few practice shots, and records his scores. Subjects are then returned to their classroom. Each is taken out of the classroom individually, led to the experimental room, and at the door is told that the experimenter is busy working on some "things" down the hall. The subject is asked to play the game alone and to bring his score sheet to the experimenter when he finishes. The door of the experimental room is then closed, and the experimenter audibly walks off.

The subject is judged to have resisted temptation if it appears he has conformed to the rules, i.e., if his score sheet shows a total score of 32 points. He is judged to have yielded to temptation if his score sheet indicates that he has altered the prearranged sequence. Transgression may be measured either by one of several highly correlated variables, e.g., the number of times a low score is skipped, the number of times a higher score than the obtained is recorded, the total score, etc., or by a weighted combination of scores, e.g., a discriminate function score.

When the subject brings his score sheet to the experimenter, he is thanked for playing the game. If his score sheet indicates conformity, regret is expressed that he has not won a badge; if it shows that the subject has

[4] The influence of peer-group status upon temptation to transgress must be investigated before this kind of incentive is employed extensively.

won a badge, it is pinned on his shirt. No attempt is made to reward those who had conformed or to relieve any possible guilt of those who had transgressed. When playing the game twice (following the procedure of the beanbag game) was tried in pretesting, it was found that the news that a badge could be won on the second try was communicated to subjects who had not yet played the game. This knowledge effectively eliminated temptation the first time they played the game.

Technical description. Figure 3 shows its schematic circuit. Scores are automatically prearranged (thus eliminating the necessity of an observer) with two 20-pole step-in relays (Ry_1-Ry_2) wired in series with Ry_8. An impulse circuit transmits current, for a few millisec., to the combination relay whenever the ray-gun trigger is pulled. The impulse circuit prevents rotation of the relay when the trigger is held down while at the same time moving the relay to a new position when it receives an impulse. The output of the combination step-in relay is fed into a program board, constructed with two 20-terminal molded bakelite strips, which controls the score light lock-in relays (Ry_4-Ry_7). The lock-in relays can be wired to the step-in relays so that at each pole any number of score lights from 0 to 5 will register. The number of lights which are lit signify the score for the immediately preceding shot. A score continues to register until the next shot is taken. The total number of shots which a subject takes is recorded with an impulse counter (C).

The ray-gun is fashioned by inserting a soldering gun trigger switch (S_1) in the handle of a pistol-type acetylene torch. The barrel of the pistol is encased in the center of a circular ¾-in. long, ⅞-in. diameter piece of relatively hard rubber which is itself inserted in a ¾-in., ⅞-in. inside diameter piece of water pipe. A 5-in. rod tapped into the water pipe is screwed into the base of the plank and provides a swivel for the ray-gun. This prevents it from being pointed upward or downward, i.e., way off target, while at the same time scores are being registered. A cord running under the plank connects the gun to the game.

The target assembly is built so that two small rockets appear a total of 64 times a minute behind the 8 in. square translucent glass. A 7¾-in. circular template, with two holes, 1 in. in diameter, drilled opposite to each other, 6 in. apart, is centered upon the shaft of a midget 4 rpm motor (M). The motor is situated so that the template rotates behind the translucent glass. A piece of black drawing paper covers the inside surface of the glass. Eight 1-in. holes, evenly spaced in a circle to correspond with the two holes in the rotating template, are cut in the drawing paper. The holes allow light to pass through whenever the holes are matched. A small facsimile of a space-rocket is glued in the center of each template hole. Red celluloid covers the drawing paper so as to provide a glowing red background for the rockets. A 6-w. bulb mounted behind each of the holes serves as a projector lamp. In order that the game will be all the more exciting, the lamps are wired to blink as the ray-gun trigger is pulled.

A push button switch (S_2) situated at the back of the game box allows the experimenter to return the combination step-in relay "home," i.e., to the beginning of the experimental sequence. Note that a cycle of 40 scores appears before the sequence repeats itself; thus, it is extremely difficult for a subject who takes several extra shots to suspect a pattern. Transgression is presumed to be facilitated by programming zero scores at the beginning and at the end of the 20-shot pattern and by following high scores with zero scores in the middle of the sequence.

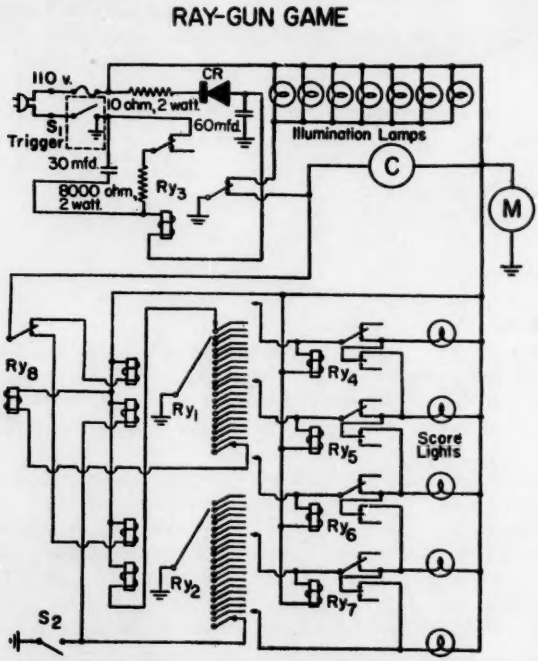

FIGURE 3—Schematic circuit of the ray-gun game. Ry_1-Ry_2 = Midget electrical reset stepper, 115 v., 20-pole. (Guardian MER 115). RY_3-Ry_7 = D.p.d.t. electrical relay, 115 v. (Potter and Brumfield GA11A). Ry_8 = S.p.d.t. electrical relay, 115 v. (one pole of P&B GA11A). S_1 = The ray-gun trigger (*see* text). S_2 = S.p.s.t. push button. C = Electric 3-digit reset counter, 115 v. (Veeder Root U-15073). M = Midget 4-rpm motor, 115 v. (Telechron). CR = Contact rectifier, 100 ma. (Federal selenium rectifier 1004A). Bulbs = 6 w., candelabra base, 115 v.

Summary

This paper has (a) set forth some of the theoretical problems in the measurement of children's temptation behavior and (b) presented a methodology which has been demonstrated to be useful in its measurement. The problems and procedures are discussed in the belief that they may be used, in either this or a modified form, for further empirical research on conflict and conscience. Temptation is conceptualized as existing for an individual when the addition or presence of a rewarding stimulus (incentive) increases the probability of a class of responses which is forbidden because of its incompatibility (conflict) with some socially expected behavior. Settling upon suitable indices of the conflict, assessing temptation to transgress, minimizing fear of detection, and holding constant goal attainability appear as some of the major problems in temptation behavior measurement. A beanbag and a ray-gun game situation have been devised, and research (3, 7) indicates that the procedures are successful. The basic methodology is essentially the same in both of the games. Subjects, asked to play the game individually, are tempted to transgress its rules in order to win an appealing incentive.

References

1. Allinsmith, W. The learning of moral standards. In D. R. Miller & G. E. Swanson, *Inner conflict and defense.* Holt, 1960. Pp. 141-176.
2. Bijou, S. W. Learning in children. *Monogr. Soc. Res. Child Develpm.,* 1959, 24, No. 5 (Whole No. 74), 25-36.
3. Burton, R. V., Maccoby, E. E., & Allinsmith, W. Antecedents of resistance to temptation in four-year-old children. *Child Develpm.,* 1961, 32, 689-710.
4. Fenichel, O. *The psychoanalytic theory of neurosis.* Norton, 1945.
5. Festinger, L. *A theory of cognitive dissonance.* Row, Peterson, 1957.
6. Freud, S. *New introductory lectures on psychoanalysis.* Norton, 1933.
7. Grinder, R. E. Parental childrearing practices, conscience, and resistance to temptation of sixth grade children. *Child Develpm.,* in press.
8. Hartshorne, H., & May, M. A. *Studies in the nature of character. Vol. I: Studies in deceit.* Macmillan, 1928.
9. Miller, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders, Vol. I.* Ronald Press, 1944.
10. Redl, F., & Wineman, D. *Children who hate.* Free Press, 1951.
11. Sears, R. R., Maccoby, E. E., & Levin, H. L. *Patterns of child rearing.* Row, Peterson, 1957.
12. Skinner, B. F. *Cumulative record.* Appleton-Century-Crofts, 1959.
13. *Oxford universal dictionary on historical principles.* (3rd Ed., Rev. with addenda) Oxford Univer. Press, 1955.

ANTECEDENTS OF RESISTANCE TO TEMPTATION IN FOUR-YEAR-OLD CHILDREN[1,2]

ROGER V. BURTON,* ELEANOR E. MACCOBY,[3] *and* WESLEY ALLINSMITH[4]

Harvard University

This study is in the tradition of many recent investigations which explore factors in the development of personality by relating childrearing practices to later behavior in the child. As part of the larger question of how one internalizes standards and rules of behavior, we ask: how does a person learn to restrain himself from doing something he very much wants to do but which his socializers have tried to teach him not to do? In this investigation, the type of resistance to temptation studied was conformity to the rules of a game. The focus will be on the relations found between childrearing practices as reported by the mother and the child's behavior during an experimental test of resistance to temptation.

Hypotheses derived from identification theories as formulated by Allinsmith (1), Maccoby (18), Sears, Maccoby, and Levin (22), and Whiting

* Laboratory of Socio-environmental Studies, National Institute of Mental Health, Bethesda 14, Maryland.

[1] We wish to express our appreciation to John W. M. Whiting whose generous help and creative suggestions have contributed to every stage of this study. Thanks are also due Robert E. Grinder, who contributed to the development of both the bean bag game and the total experimental design, and to the staff of the Laboratory of Human Development, Harvard University, who aided in gathering the data. We also thank Drs. John D. Campbell, John A. Clausen, Thomas L. Gillette, and Marian Radke Yarrow for constructive comments on early drafts of this paper.

[2] The main support for this study was provided by a grant from the National Institute of Mental Health (M-1096).

[3] Stanford University.

[4] University of Cincinnati.

Child Develpm., 1961, 32, 689-710.

(27, 29) mainly determined the selection of childrearing practices to measure. These formulations, taking much from psychoanalytic writings, propose that the child adheres to parental standards as a consequence of an identification process whereby the child patterns himself upon a parent. Two motives are generally posited for this process: the desire to reproduce the gratification provided by a loved object and the desire to control, neutralize, or protect oneself from a fearful, threatening, or powerful object. Whiting (29) suggests that these two kinds of identification are really based on the single motive of envy of the controller of resources.

Another body of psychological theory based on the effects of direct reward and punishment provides parallel and sometimes different hypotheses concerning origins of resistance to temptation. According to this general behavior theory a concomitant of punishment is fear which becomes associated with the stimuli present when parental punishment of a forbidden response is administered. This fear of punishment aroused by the temptation to perform the forbidden behavior then acts as an acquired drive (20, 26) or conditioned negative reinforcer (7, 21, 25) to deter a child from performing the response even when the parent is not present to punish directly.

Much of the hypothetical formulation which guided the study grew out of findings of previous investigations in which childrearing practices and measures of guilt were associated (1, 2, 3, 8, 19, 22, 30). Since it was assumed that a child who had experienced guilt would later tend to resist temptation in order to avoid the unpleasantness of remorse, it was expected that the antecedents for guilt, found in these previous studies, and those for resistance to temptation would tend to be the same. These earlier findings (as well as other theoretical considerations) led to the expectation that severity of weaning and toilet training (30) and certain aspects of sex training (11, 24) should be related to resistance to temptation. It was expected, too, that associations found earlier of guilt with parental use of reasoning and love-oriented techniques of discipline (in contrast to arbitrary and physical punishment) would also be found between resistance to temptation and these parental disciplinary practices (1, 3, 8, 19, 22, 30). In addition, our theory of identification led us to expect that children of restrictive parents (as contrasted with permissive parents) would be more likely to engage in extensive covert role-practice of their parents' actions and statements, and thus would more thoroughly "internalize" parental values and be more likely to resist temptation.

Without reference to the concept of identification, it is possible to make predictions concerning the features of socialization which ought to be related to resistance to temptation, simply on the basis of reinforcement learning theory. First of all, consistency should be important. If the parents always interpret the same acts as being "wrong" and they employ methods (such as keeping close track of the child's whereabouts) for insuring that the child is almost always caught when deviating, the child's resistance to temp-

tation should be increased. Presumably, both fairly high levels of punishment for deviation and reward for conformity should also be associated with high resistance to temptation. Furthermore, Solomon and Whiting[5] have suggested that temporal parameters involved in the application of punishment for deviations may be crucial in producing resistance to temptation versus guilt. According to this theory, the stimuli which have been present during the administration of punishment become, through conditioning, capable of arousing anxiety in anticipation of the actual punishment. If the parent has usually administered punishment *before* the child could commit the forbidden response, the cues producing anxiety should be maximum when the child is approaching the tabued acts, i.e., is being tempted, and thus should produce maximum resistance to temptation. On the other hand, punishment usually administered *after* the response has occurred should condition anxiety to the cues present at that time and thus be related to guilt or feelings of remorse after a deviation. Whiting's specific resistance to temptation hypothesis (27) adds to these conditions the prediction that punishment administered by a generally nurturant agent is more effective in producing resistance. There is some evidence from studies with dogs which partially support these hypotheses (6, 9, 27).

Some previous investigations related conditions of the home background with conformity to rules but were either very general (14) or obtained the background measures from the subject's own report (19). More recent research by Sears, Maccoby, and Levin (22) related many areas of child-rearing practices to a rating of the child's current evidence of conscience, all measures based on mother interview data. The present study replicates much of this latter investigation and also provides a behavioral measure on the child to be used as an index of resistance to temptation.

In line with our theoretical interests and the final research design, 15 general areas of parent-child characteristics relevant to childrearing practices were considered: activity level of the child now and during infancy; weaning; toilet training; sex training; aggression training; *S*'s tendency to be emotionally upset; restrictions in infancy and currently; independence training; level of demands; clarity and consistency of demands, rules, and restrictions; mother's attitude toward *S*'s knowledge of games, rules, and cheating; techniques of discipline; use of positive reinforcement; mother's affectional warmth; and father's affectional warmth and level of demands. The general questions which this report considers in relation to these content areas are:

1. Do early childrearing practices—specifically, weaning, toilet training, and sex training—relate to conformity to rules in the four-year-old?

2. Are techniques of punishment (love-oriented as contrasted with physical punishment or object-oriented discipline) related to resistance to temptation?

---

[5] Personal communication, 1956.

3. Does the closeness of parental control relate to resistance to temptation? Two facets of this question deserve study: (a) parental consistency and (b) the general level of demands and restrictions.

4. Is resistance to temptation related to the timing of punishment (that is, whether punishment is usually administered just before, instead of after, the deviant act)?

5. Is resistance to temptation related to the identity of the disciplinary agent or to the degree of affectional warmth that characterizes this agent?

Method

Subjects

The 77 children, 40 boys and 37 girls, in this study were all four-year-olds and enrolled in private nursery schools in four districts of the greater Boston area.[6] The *Ss* were tested between April and June of 1957, and the mothers were interviewed between May and August. In general, these *Ss* came from well-established, middle-class homes. Only six fathers and 18 mothers had not attended college. Most of the fathers were business or professional men. The religious preferences were Jewish for 38 and Protestant for 30 of the families with the other nine families being mainly mixed in preference. The ethnic backgrounds of the parents were chiefly British, Northwestern, or Eastern European.

Procedure

The major dependent variable is a behavioral measure of whether or not the child conformed to rules of a game when he played it by himself. A bean bag game was built which was interesting to the *Ss*, appeared to be controlled by the player's skill, permitted deviation from the rules easily, and yet was in fact completely controlled by an experimenter. The game consists of a 1 ft. by 4 ft. board on which there are five lights in a row. As presented to the *S* and as it appeared to the *S*, hitting a wire which was stretched along behind the board (but which he could not see from his throwing position) would turn on a light and ring a bell. The game was placed against a wall to make it seem plausible to the child that if a bag hit high on the wall it could fall down on the wire and make a score. Regardless of how many times *S* hit the wire with his bean bag, however, he got a score only when a hidden experimenter turned on a light. Individual differences in skill were thus controlled, as were schedules of reinforcements.

In order to test the children in a rather familiar setting as well as to avoid the costs of transporting them to the university laboratory, a portable

[6] Originally 85 *Ss* were tested, but only 79 of the mothers were located for home interviews. Two of the obtained mother interviews were discarded due to failure of recording equipment.

observation booth was built. This booth and the game were installed in an empty room of the *Ss*' nursery school.

Each *S* was shown how the game "worked," was given standardized practice in throwing the bags, and was carefully instructed in the rules. The rules were to throw each of the five bags only once over the board while standing on a foot marker placed approximately 5 feet from the front of this board.

After it was clear that *S* understood the rules, a second game was introduced.[7] After three minutes of the second game, *E* brought *S* back to the bean bag game and reviewed the rules. Then, a tray of toys, pretested for their attractiveness, was shown *S*, and he was asked to select one he would take were he to "get enough lights on to win a prize." This was done to be sure he focused on them enough to develop a strong desire to have one.

E then looked at his watch and said he had to make a phone call but that he wanted *S* to play the game according to the rules while he was out of the room. The implication was that *S* was to play for the prize. *S* was shown how to hook the door from the inside, so that no one would "bother" him while he was playing the game. *E* said he would knock when he wanted to be readmitted.

If *S* played according to the rules during the three-minute test, he received only one light (after his second throw) in contrast to the three lights "earned" in each practice session. Additional lights were given for each act of cheating: stepping forward, moving the foot marker, retrieving bags incorrectly, and hitting the wire with the hand. During this time, the hidden *E*, controlling the game, recorded *S*'s behavior.

When *E* returned, another game was played, "this time for the prize." This statement was intended to reduce any guilt that might have been aroused by cheating, to relieve feelings of failure for noncheating *Ss*, to avoid reinforcing cheating, and to have a check on whether *S* evinced understanding of the rules by playing according to them with *E* present. All *Ss* were given four lights in this last game and awarded a prize.

Before continuing, we would like to explain part of this procedure. It can be argued that many motives are aroused by the temptation test. Examples of motives that would lead to obtaining a high score through cheating are (a) to please *E* by throwing well, (b) to fulfill *S*'s own achivement strivings, competitive tendencies, and/or need to outwit authority, and (c) to win the toy. Motives that would promote resistance to temptation include (d) to please *E* by conforming to his rules, (e) to avoid external punishment (actual or imagined), and (f) to comply with *S*'s own internalized standards. Instead of attempting to control each possibility of these sorts, our decision was to test resistance to temptation under conditions which

[7] Part of the full procedure irrelevant to this presentation will not be discussed here. The complete description is in Burton, Allinsmith, and Maccoby, "Resistance to temptation in relation to sex of child, sex of experimenter, and interruption of attention" (not yet published), which presents material on the effects of the experimental variables.

attempted to maximize all possible motives to deviate. Thus, intensity of temptation in toto would be controlled, and individual differences in temptation would be less readily attributable to mere variations in desire to transgress. To minimize risk and to maximize temptation, all windows through which *S* could be observed were covered with curtains, *E* acted pleased when *S* was given a "hit" during the practice games, the toys were shown and *S* was asked to choose the one most attractive to him, *S* was instructed to hook the door from the inside during the test, and each *S* was given the same reinforcement schedule of only one "hit" during the test if he conformed to the rules. In learning the rules, *S*s got three "hits" in each practice game. Thus, we felt confident that they considered getting just one light was a poor showing.

An interview schedule was developed to obtain information for rating those childrearing variables and indexes of early childhood behavior considered theoretically important. Most of the questions were open-ended, and the interviewers were trained to make appropriate probes in order to get sufficient details for scoring. All interviews were recorded on Audograph discs, and transcripts made later. The interview averaged two hours and was very much like that described in Sears, Maccoby, and Levin (22). There were, however, some variations in content.

Parent training measures. We sought indexes involving as few inferences as possible. Therefore, ratings scored in each of the content areas were generally based on specific information in the protocol, rather than on global judgments.

All interviews were scored by the senior author. Forty-four interviews were coded by at least two people; but, since some interviews had been used for training, only 24 were used for testing reliability. Of the 142 scales scored, 28 were discarded for not reaching a Pearsonian r of .70. The remaining scales had reliability coefficients ranging from .70 to 1.00, with a mean of .86.

Summary scales were made of the childrearing variables on the reasoning that each original scale could be considered analogous to an item in a test of a general area of child training. The method of computing these summary scales was first to choose the individual scales considered psychologically related, that is, as potentially measuring the same underlying training area. Then, tables of intercorrelations were made, and from these tables those scales found to have internal consistency were chosen. The criterion for internal consistency was that any rating to be included must have a majority of positive correlations with the other ratings and should not have any significant negative correlations with any of the ratings finally chosen. The scores of the original scales were then added. These sums were categorized into 5-point scales as close to normal distributions as possible for the final summary score.

Child measures. The resistance to temptation score was based on the child's conformity to the rules of the game during the temptation test. This

was a 7-point scale based on a count of the number of bags *S* threw correctly before cheating, with categories at the extremes for immediate cheating and for no cheating during the test. Thus, "1" was for deviating before any bags were thrown; "2" through "6" were the scores obtained by adding 1 to the number of bags correctly thrown before first deviation; and "7" was for never deviating during the test period. This measure had almost perfect reliability.

An "evidence of conscience" measure was obtained from a rating of the mother's report in the interview, the same measure as in Sears, Maccoby, and Levin (22). This measure was made from the mother's report of how *S* behaved when he had done something wrong and before the mother knew about it. This rating had a scoring reliability coefficient of .83.

Results

The individual measures were divided at the median and fourfold tests of association made. Whenever possible, Fisher's exact probability tests were used; otherwise, χ^2 was computed, with all probabilities based on two-tail tests and corrected for continuity. The summary scales were correlated with both child measures using product-moment correlation.

Childrearing and Early Childhood Behavior Related to Resistance to Temptation

Table 1 summarizes the relations between the individual childrearing scales and the resistance to temptation measure. Also included are a few indexes of child reactions to socialization practices: the information about these reactions was also obtained in the mother interviews. Only those scales significantly associated with resistance to temptation or of special theoretical relevance are included.

Activity level. The three activity scales were highly intercorrelated, with correlations decreasing in magnitude as the age differences increased. Both infancy (before one year) and early childhood (between first and second year) activity levels were positively related to resistance to temptation for boys (with a corresponding but nonsignificant trend for girls). The current activity (age at the time of interview) rating, however, was not significantly related to conformity to rules, and the data are not presented.

Early childrearing. The only weaning variable which was related to temptation was an over-all rating of severity which was positively related (for boys only). The more severe the weaning, the more conformity to rules. This rating measured how gently the mother changed *S* from sucking to drinking, at the low end of the scales, and how abruptly and severely, at the other end. Age at beginning and completion of weaning were separate ratings and proved to be uncorrelated with resistance to temptation.

Time between initiation and completion of bowel training and the age when *S* stayed dry during the day were related to resistance, but again these

relations were found only among boys: the longer bowel training took and the later the completion of daytime dryness training, the more conformity to rules. The daytime dryness variable was dichotomized at the median for each sex before being combined for the test of association. The median for girls was 24 months, for boys, 34 months. While it was also found that the age for beginning bowel training was significantly related to the amount of time it took for completion (the earlier training began, the longer it

TABLE 1

CHILDREARING AND CHILD BEHAVIOR RELATED TO RESISTANCE TO TEMPTATION

		RESISTANCE TO TEMPTATION				
		Boys		Girls		
		High	*Low*	*High*	*Low*	*Significance*
Activity level during first year	High	14	4	13	9	Boys* Girls Combined: $\chi^2=3.67$
	Low	8	13	7	7	
Activity level between one and two years	High	14	4	11	5	Boys* Girls Combined: $\chi^2=7.53^{**}$
	Low	8	14	9	12	
Severity of weaning (age of beginning not considered here)	High	14	5	11	9	Boys* Girls Combined: $\chi^2=2.55$
	Low	7	12	8	8	
Time between initiation and completion of bowel training	"Short"	5	13	7	7	Boys** Girls Combined: $\chi^2=4.00^{*}$
	"Long"	13	4	12	10	
Age when stayed dry during day	Early	8	15	7	10	Boys** Girls Combined: $\chi^2=7.85^{**}$
	Late	13	3	13	7	
Explanations given about origin of babies						
Rarely		10	15	13	8	Boys* Girls
Often (whenever *S* might ask)		11	3	7	9	
S's reaction when he learned about sex differences	High	6	9	3	9	Boys Girls* Combined: $\chi^2=6.46^{*}$
	Low	16	9	16	7	
S's reaction to being sent to room	High	9	10	8	7	
	Low	13	7	12	9	
Timidity	High	7	5	9	7	
	Moderate	7	9	6	4	
	Low	8	4	5	6	
Techniques used for teaching S not to touch objects						
Stops before touches		9	12	7	11	Boys Girls Combined: $\chi^2=4.37^{*}$
Waits to see if he really plays with it		12	6	13	6	

(*table continued on next page*)

TABLE 1 (*continued*)

CHILDREARING AND CHILD BEHAVIOR RELATED TO RESISTANCE TO TEMPTATION

		RESISTANCE TO TEMPTATION				
		Boys		Girls		
		High	*Low*	*High*	*Low*	*Significance*
Mother's attitude toward cleanliness	Relaxed	13	2	4	5	Boys** Girls
	Concerned	9	16	16	12	
Does S understand what cheating means?	Yes	10	7	5	9	
	No	12	11	15	7	
Usual technique of discipline						
Reasoning		0	0	1	1	
Visible suffering		0	0	0	1	
Withdrawal of love		0	0	1	0	
Isolation		7	8	6	5	
Deprivation of privilege or object		3	4	0	4	
Scolding		2	2	7	3	
Physical punishment		10	4	4	3	
Scolding and physical punishment vs. all others: Boys Girls Combined: $\chi^2=2.79$ ($p<.10$)						
Frequency of Mother's use of reasoning as teaching technique	High	8	12	5	9	Boys Girls Combined: $\chi^2=5.41$*
	Low	14	6	15	8	
Use of "I'm not going to talk to you until you behave better"	Yes	12	3	5	6	Boys* Girls
	No	10	15	15	11	
Use of deprivation of privileges	High	10	8	8	13	Boys Girls*
	Low	10	10	12	3	
Frequency of Mother's playing with S just for her own pleasure	High	14	8	10	15	Boys Girls*
	Low	8	10	9	2	
S's initial anxiety in the experiment	High	5	9	8	1	Boys Girls*
	Moderate	8	7	8	9	
	Low	8	3	4	7	

NOTE.—Throughout this table the variable being considered was divided at the median.
* $p \leq .05$. ** $p \leq .01$.

took), it is nevertheless true that the age of beginning bowel training was not in itself related to resistance to temptation. Furthermore, our rating on severity of toilet training did not predict resistance to temptation.

The only rating of sex training to be significantly related to resistance to temptation concerned explanations about origin of babies. Parental offering of such explanations in response to a child's query was associated with resistance for boys. The trend is opposite for girls, though nonsignificant. Although we had reliable scales on the degree of sex permissiveness and degree of mother's sex anxiety, they did not correlate with the children's

resistance to temptation. The original reason for including these measures of sex training was based on the hypothesis that children with high sex anxiety, whatever the contributing factors, would tend toward an inability to delay gratification which in turn would incline the child to cheat in the temptation test. In general, then, the conditions which ought to produce sex anxiety in the child—namely, severe, nonpermissive training in regard to sex—did not prove to be associated with yielding to temptation, except for the finding about giving information to boys. Another significant association interpreted as supporting this hypothesis is that between cheating and a high reaction when *S* first found out about sex differences (especially for girls), such as a great amount of curiosity or feeling that something was wrong with himself or herself.

Anxiety. In addition to anxiety about sex, it was hypothesized that in general a highly "anxious" child would be low on self-control and delay of gratification with an increased tendency to cheat. The indexes of anxiety in *S* based on the interview were ratings of his general timidity in meeting people, amount of upset over making mistakes, regression to wanting a bottle or to wetting the bed, amount of upset when sent to his room, as well as intensity of reaction to sex differences. The only "anxiety" rating from the interview which was related to resistance to temptation was the high reaction to sex differences associated with cheating. Table 1 shows the distributions on the interview ratings of *S*'s reaction to isolation and to meeting strange people. The other ratings of *S*'s emotionality were evenly distributed. In addition to the ratings based on the interview, there were two ratings of the child's "anxiety" during the experimental situation. An "initial anxiety" rating was made from *S*'s behavior upon being brought into the experimental room and was based on the amount of hesitancy he showed when leaving the classroom, fearfulness of *E* and the bean bag game, and reluctance to permit the teacher to leave him when coming into the experimental room. The other "anxiety" rating was based on *S*'s reaction to *E*'s withdrawal of attention during the experimental situation. The latter showed no trends, but the relation between the behavioral rating of initial anxiety and resistance to temptation is shown in Table 1. It is seen that for boys there is some tendency in the distributions for a high "anxiety" rating to be associated with cheating. The distributions for girls suggest that two types of "anxiety" have been measured. One type seems to reflect the emotional instability which it was originally hypothesized would be related to cheating; the other type suggests a fearfulness of responding. These findings will be discussed further below.

Measures of aggression training had very low reliabilities and therefore did not permit a test of the relationship between this area of child training and cheating.

Timing of punishment. It was originally thought that the measure of the parent's timing in trying to stop *S* from touching something would be related to the ratings on restrictions, but it was not. This scale was, however,

related to resistance to temptation. The categories for the scale ranged from stopping *S* before he touched the object ("1") to permitting him to touch it and then telling him not to touch it any more ("5"). The association indicated that letting a child actually touch objects before saying "no, no" is related to resistance.

Parental control. Only one of the 25 ratings measuring restrictions, demands, and autonomy was significantly related to resistance. This rating was the mother's attitude toward cleanliness, and the association was only for boys; the greater the mother's concern about cleanliness, the greater her son's tendency to cheat.

None of the scales attempting to measure the mother's consistency and clarity of rules, demands, and restrictions discriminated between high and low resistance. This was also true for the ratings of the mother's attitude about rules. For these last-named ratings, the mother was asked how often *S* played games with rules, whether he understood about rules and cheating, and how often she consciously taught him about cheating. It seems of interest to note that a mother's opinion of her son's knowledge of rules had little relation to his behavior; in contrast, girls whose mothers felt they definitely understood about following rules and about cheating tended to cheat. This tendency for girls which is nonsignificant when the Resistance to Temptation (RTT) scale is divided at the median becomes significant if "noncheaters" are compared with "cheaters" (Table 2).

TABLE 2

UNDERSTANDING OF CHEATING RELATED TO RESISTANCE TO TEMPTATION

	BOYS		GIRLS**	
	Noncheaters	*Cheaters*	*Noncheaters*	*Cheaters*
Does S understand about cheating?				
Yes	7	10	0	14
No	8	15	12	10

** $p \leq .01$.

Techniques of punishment. In regard to disciplinary techniques, physical punishment included such acts as spanking, slapping and shaking. Scolding was verbal reprimand, usually with a raised voice. The mother's taking away a toy or TV was scored deprivation of object or privilege. Isolation was primarily scored for sending *S* to his room. The prototypical response for withdrawal of love would be, "I can't love you when you do that." Visible suffering (4) was represented by the mother's saying to *S*, "You must not love me to act that way." Some mothers claimed to discipline their children by discussing why they wanted them to act in certain ways and

expressed the attitude that they didn't believe in "punishment." The major disciplinary technique for these mothers was "reasoning." A rating of the mother's usual choice of discipline indicated a suggestive trend that scolding and physical punishment tend to be related to conformity when constrasted with all other disciplinary techniques (p between .10 and .05).[8] The other measures of physical punishment and scolding, such as frequency and intensity, did not show any strong trends of association with resistance. A consistent trend was found for high use of reasoning to be related to cheating. This tendency was found in the scales comparing preferred types of discipline and mother's belief in reasoning and was significant in the measure of the extent the mother used reasoning. For withdrawal-of-love techniques, only high use of not talking to *S* until he behaved was related to resistance, but only for boys. The only other negatively centered training technique which showed a significant association with temptation behavior was the high use of deprivation of privileges which was related to cheating for girls only.

The measures for the positively focused training techniques did not show any significant associations with behavior under temptation (and are not shown here), but the trends in the distributions were rather consistent. For boys the trend was for high use of positive reinforcement, such as praise, affection, tangible rewards, and privileges, to be related to resistance. The pattern for girls was for the use of privileges and tangible rewards, labeled "object-oriented techniques" by Sears, Maccoby, and Levin (22), to be related to resistance, but for high use of love-oriented practices of praise and affectional reinforcement to go with cheating.

Identity and warmth of disciplinary agent. The measures of the mother's warmth asked about her attitude toward *S*, what she liked about him, how demonstrative she was toward him in showing affection, how often she played with him for her own pleasure, and whether she took time to answer his questions. The only significant relation for these measures was for girls: those whose mothers said they frequently played with their daughters just for pleasure tended to cheat. For boys, the consistent trend, though not significant statistically, was for warm mothers to have sons who tended not to cheat. The same nonsignificant trends for girls and boys were found in the measures of father warmth; high warmth tended to be related to cheating in girls and conformity in boys.

Resistance to Temptation and Evidence of Conscience (Guilt)

In most theoretical formulations based on the concept of identification, it has usually been assumed that resistance and guilt would be positively

[8] The senior author made an over-all rating of the mother's preferred mode of discipline —psychological, reasoning, scolding, or corporal—which indicated corporal punishment was related to conformity for boys and girls. However, a check showed the rating to be too unreliable for any confidence. We would like to express our appreciation to Mrs. Faye Goldberg and Dr. John D. Campbell for performing this reliability check.

related. The reward and punishment theory, however, led to the hypothesis that conditions for producing resistance to temptation and for producing guilt might be different, namely, administration of punishment before the act versus after the act, respectively. A test of association between these two "superego" indexes, resistance to temptation in the game and evidence of conscience, indicates that there tends to be a negative correlation between them. If all "cheaters" are compared with *Ss* who never deviated from the rules during the entire three-minute test period, this negative association is less than the .10 level of probability.

TABLE 3

EVIDENCE OF CONSCIENCE RELATED TO RESISTANCE TO TEMPTATION

	BOYS		GIRLS	
	Noncheaters	*Cheaters*	*Noncheaters*	*Cheaters*
Evidence of conscience (Guilt)				
High	8	18	4	16
Low	7	7	8	9
	Sexes combined: $\chi^2 = 3.12$ ($p < .10$)			

Childrearing and Evidence of Conscience (Guilt)

Sears, Maccoby, and Levin (22) have presented a theoretical formulation for the development of conscience. They reported correlations between childrearing practices and evidence of conscience (guilt) in the child which, although small and not all significant, were in support of their theory. As a check on the stability of these findings, the identical measure of conscience was obtained in this study and has been correlated with the summary scales for the childrearing measures. The resulting correlations were at least of the same magnitude as those previously reported and were generally higher. Low use of physical punishment (−.30),[9] high use of reasoning (.18), high use of praise as reward (.25), and high maternal warmth (i.e., affectional demonstrativeness) (.22) were significantly correlated with a rating of high conscience in both studies. Also considered theoretically important but not found to be significantly correlated in the previous Sears *et al.* study, the following practices were significantly associated in the present investigation: low use of tangible rewards and deprivation of privileges (−.16) and high use of isolation (.15).

In the present study the mother was asked several questions about her attitude toward the child's knowing about and following rules. The sum-

[9] Numbers in parentheses are product-moment correlations between the summary scale and the conscience rating in this study and are significant at or beyond the .05 level.

mary scale for these ratings was correlated .34 with the conscience rating, the highest of all the summary scales. This positive correlation indicates that the mothers who felt it was important for their child to learn about and to observe rules also felt their child evidenced many signs of a well-developed conscience. This somewhat severe attitude regarding conformity to rules being related to conscience contrasts to some extent with the "permissive" attitude expressed by the other correlations.[10] Considered together, however, these relations do suggest an internal consistency of childrearing beliefs shared by the mothers of this study. This consideration becomes more cogent in the light of the general tendency of these scales to correlate in the opposite direction (although not significantly) with our measure of resistance to temptation. This will be discussed further.

Discussion

An unexpected finding was the positive relation between the child's activity level (as reported by his mother) and his ability to resist temptation. We can only speculate about the reasons for this relationship. Two possibilities come to mind: the first is that active children are those in whom the parents have *permitted* activity, whose parents have not emphasized being quiet but have allowed the child to explore the environment fully. Perhaps, children treated in this way are simply more mature in all respects, including the kind of ego organization that makes possible conformity to rules. But this explanation does not seem entirely consistent with the fact that it was the child's level of activity *up to the age of two*, not after, that predicted resistance to temptation. We suggest an alternative explanation: parents who consider their child more active, either because the child is innately more active or because of the parents' own attitude toward how active a child should be, will feel that their child more often exposes himself to dangerous or forbidden things; these parents are therefore more often stimulated (by the actions of the child) into controlling actions designed to protect the child or the family's property; and, therefore, children who are considered highly active during the reaching-and-touching stage receive more training in *not* touching, a condition which facilitates the child's later ability to resist temptation in the absence of a controlling agent.

Our findings on the relation between early childhood socialization practices in the areas of weaning, toilet training, and sex training are fragmentary and not fully consistent. Severe weaning (usually meaning abrupt weaning in which the mother refuses to comply with the child's demands for the return of the bottle) was associated with high resistance to temptation in boys, as was *long-continued* bowel training, again for boys only. In the latter case it is notable that an early start of bowel training in itself was not significantly associated with resistance to temptation, nor was late com-

[10] In the present study low severity of toilet training (−.22) and low activity (−.16) were also correlated with high conscience.

pletion of training. What is the special meaning of long-continued training? Often it means that a mother begins relatively early but for some reason does not succeed. Long-continued training frequently means that toilet training has become a battleground between the mother and child. The mother applies her demands unskillfully, with poor timing, and the child goes through periods of negativism and regressive behavior. Long-continued training appears to mean a greater total amount of pressure in this behavior-system than brief training does, be it early or late. We may have some basis, then, for the generalization that heavy socialization pressure on boys (as exemplified by severe weaning and extended toilet training) leads to greater resistance to temptation at a later time. But this interpretation must be tentative, considering that our scale on *severity* of toilet training did not correlate with resistance to temptation, nor did the amount of socialization pressure applied by the mother in the area of sex training.

For boys the indexes of anxiety based both on the interview and on the behavioral ratings during the temptation test consistently showed trends which suggest that a high emotional reactivity is related to cheating. In girls there seemed to be two kinds of anxiety measured. One type, indicated by fearfulness of new situations and by timidity when meeting new people, tended to be associated with noncheating. This kind of anxiety seemed to predispose the child to feelings of fear during the test. The other type of anxiety seems to reflect a more general emotional instability and inability to delay gratification. This form of anxiety was suggested by the measures of *S*'s reaction to sex differences, to isolation, and to withdrawal of attention, and was associated with cheating. However, these interpretations are rather speculative since they are based as much on trends as on significant findings.[11]

In our introductory section, we discussed the possible outcomes of different timing of the mother's socializing actions. We mentioned the hypothesis that stopping a child *before* the execution of the forbidden act might lead to resistance to temptation, while stopping him after the action has actually begun might lead to guilt over deviation. Our finding does not confirm this prediction. The question we used in the mother interview was:

> Suppose you wanted him to learn not to play with something—like the TV set or matches. If you saw that he was tempted to touch it, would you stop him before he touched it or wait to see if he really played with it and *then* correct him?

The answers to this question do not provide us with exact information concerning the mother's timing of discipline. "Stopping before he touches it" covers a range from diverting the child to some other activity before he

[11] These two hypothesized kinds of anxiety seem to be similar to the factors found by Bell (5) to discriminate between neonates; the "tactile sensitivity" factor which is based essentially on the baby's reaction to aversive stimuli is analogous to fearfulness of new places, strange people, pain, etc.; the "arousal" factor seems similar to what we have considered emotional upset or instability.

gets very close to the forbidden object to waiting until he actually starts putting out his hand and slapping the hand in the split-second before it touches the object. And similarly, "waiting till he really plays with it" could include applying discipline in the moment just following the child's first touching the object to letting him play with it for an extended period before stopping him and scolding or punishing him. We suspect that, since the objects asked about are either dangerous (matches) or quite expensive to repair if damaged (TV set), very few of these mothers actually allowed the child to play with the object to satiation. We suspect that the meaning of "wait till he actually plays with it" was more a matter of the mother's wanting to make sure that her intuition about his intentions was correct before applying discipline and that the discipline usually followed fairly closely upon the onset of actual touching. If this interpretation is correct, our finding becomes reasonable from the standpoint of learning theory, for it means that the correlation between discipline and the relevant cues for this particular action was very close in the "wait till he touches" condition. If the mother applies discipline when she only *thinks* the child is about to deviate but does not wait to make sure, she will sometimes be right and sometimes wrong, and the result will be poor correlation between the cues the child is actually responding to and the consequences of the action—an inefficient condition for learning. It is clear, however, that we need data on the actual timing of positive and negative reinforcement, recording much finer gradations in time intervals than can be obtained from interview questions. We expect theoretically that applying the socializing action *immediately* before or during the first phase of the subject's forbidden action would be more effective than applying it either earlier or later. But our data do not permit us to determine whether this is so.

The extent and severity of rules, restrictions, and demands, and the mother's consistency and clarity in establishing them would seem to be extremely important in establishing standards of behavior in *S*. However, the ratings which attempted to measure these variables showed few relations with response to temptation. The dearth of relations does not mean these variables have no effect on resistance; but it would seem to suggest that these variables, if indeed they *are* important, may have to be measured by tools other than the interview. Or perhaps the difficulty lies in our definition of what constitutes demands and restrictions.

It is difficult to account for the finding that girls who cheated tended to have mothers who felt their daughters understood about rules and cheating, but boys whose mothers said they understood about rules tended slightly to conform to them during the test. It may be there is a difference in the mother's attitude about what is referred to by "rules of the game" depending on the sex of her child. Games involving physical power and skill (games similar to our bean bag game) may be her referent for boys, whereas games more unlike our temptation test are associated with games played by girls. It is also possible that when mothers decide to teach their

children to conform to rules they are more effective teachers for their sons than with their daughters, a possibility suggested by the Gewirtz and Baer (10) findings that social reinforcement is more effective when administered by an opposite-sexed experimenter.[12]

The indication that physical punishment has a tendency to be associated with resistance as compared with psychological punishment and use of reasoning would seem to be at odds with the implications from findings in previous reports (2, 3, 19, 22). Although the association between these variables reached a significance level just slightly under .10, the trend seems especially noteworthy because of its theoretical significance and its marked opposition to the expectation from these previous conclusions. These discrepancies may be due to poor measurement, differences in dependent measures, differences in ages of *Ss*, and/or sources of measures. In Allinsmith's (2, 3) studies the measures of guilt and resistance to temptation were based on story completions and employed older *Ss* than used in this study. MacKinnon (19) obtained his measures of parental discipline from a retrospective account of the *Ss* who were also older. Sears, Maccoby, and Levin (22) had the same age *Ss* but based both antecedent and consequent measures on the mother interviews, with the consequent measure being of guilt rather than of resistance to temptation. In contrast to the conclusions of these previous investigations, the trends in the present data suggest that scolding and physical punishment may be more efficient in producing resistance to temptation in four-year-olds than is psychological punishment, and certainly more effectual than attempting to train by just "reasoning." This evidence also suggests that, at least at this age, just understanding what an adult wants him to do or fearing psychological punishment is not as effective in deterring deviation by a child as is fearing physical punishment.

Our results then suggest that there is no simple relation between "permissive" or "severe" childrearing practices and resistance to temptation. Some training techniques which we found associated with conformity to the rules seem to reflect severe socialization practices. This severity grouping is suggested by the associations between the behavioral conformity and the interview measures of high activity, severity of weaning, long toilet training, use of scolding and physical punishment, and low use of "reasoning" as a disciplinary technique. On the other hand, resistance was also associated, for boys only, with explanations about sex, permissive attitude toward cleanliness, and not talking with him when he misbehaved (psychological punishment), and, for both boys and girls, with permission to touch an object before saying "no." This latter "cluster" of training practices suggests an emphasis on appealing to the child's cognitive awareness of what the parent wants. These groupings indicate the possibility that the differ-

[12] For a somewhat different conclusion from other data, see Gump and Kounin (13). These authors conclude that boys are behind girls in the development of school-related value orientations due to developmental lag, identification difficulty, and/or intensification of preschool sex differences. This interpretation seems to correspond to that of Sears, Maccoby, and Levin (22) that boys lagged behind girls in the development of conscience.

ences in results between this study and the previous studies, regarding psychological and physical punishment, may in large part be due to the differences in the ages of subjects: previous studies have used older *Ss*. The hypothesis is suggested, here, that the more direct, severe, physical modes of training are most efficient and effective at the early ages of development. But with greater cognitive development, and perhaps increasing identification, the psychological techniques gain in power to influence the child's internalization of parental standards. The argument is that the direct forms of training represent a procedure similar to conditioning employing an unconditionally aversive stimulus; whereas at a later age the child is able to "comprehend" the more symbolic, psychological forms of punishment which may have acquired their aversive value.[13] This hypothesis of cognitive development also might account for the ineffectiveness of "reasoning" as a disciplinary technique for these four-year-olds.[14]

The trend that boys from homes with high parental warmth tended to resist temptation, although not significant, is consonant with the results of the previous study (22). The direction of this result also supports Whiting's hypothesis that the administration of sanctions by a normally nurturant agent will be more effective (27). In contrast, a "cold" home was associated with conformity to rules for the girls. Sears, Whiting, Nowlis, and Sears (23) and Hartup (15) have advanced an identification hypothesis to account for this differential effect. Their argument is that girls are more sensitive to punishment by the mother because the mother is the girl's identificand. The finding which shows high use of withdrawal of love by the mother related to conformity for boys and not for girls in the present study tends to be contrary to this hypothesis. However, the findings that mothers in general are more nurturant to girls, especially during infancy, and use "psychological" forms of punishment more frequently with girls and physical punishment more frequently with boys (22, 23) are background conditions for considering these two hypotheses and indicate that they are not in opposition.[15,16]

---

[13] It seems plausible that psychological discipline may become even more "painful" to *S* than corporal methods once he learns to respond to it, in that, if the parent slaps him, the punishment is over; but the mother's not talking to him for half an hour lasts longer.

[14] An investigation by Robert E. Grinder using 11-year-old subjects gives some support to this hypothesis (12).

[15] It has been reported that mothers of girls "reverse their policies, in terms of severity of frustration, from the infancy years to the later preschool years" (i.e., from high warmth and nurturance to a relatively greater severity of demands) and are "more consistent with respect to this variable" for boys (23). This evidence would suggest that the currently "cold" mothers of girls may have been "warmer" during the infancy years—a condition which brings together the hypothesis regarding the greater effectiveness of a basically nurturant agent with the identification hypothesis for our data.

[16] It is possible that innate constitutional differences between the sexes operate in regard to sensitivity to physical punishment. There is some evidence that the pain receptors of girls have a lower threshold than those for boys (5, 16). This greater sensitivity to pain would produce more fear associated with the cues of punishment and should lead to increased resistance to temptation for girls relative to boys under equal amounts of physical punishment. Again, this hypothesis should be considered in light of the evidence that boys

The negative relation between the measure of resistance to temptation and the rating of conscience lends support to the hypothesis that the conditions for guilt and resistance to temptation are different. This hypothesis that moral development is a multiple rather than a single process was offered by Allinsmith (1, 2) as an interpretation of his findings and has been discussed by Maccoby (17) in a paper which utilizes additional data from recent and still unpublished studies. It is also possible that a greater number of significant relations which we found between the childrearing variables and the conscience measure, as compared to the temptation measure, indicates a "halo" effect arising from the fact that the conscience measure and the childrearing variables are both taken from the mother interview. Furthermore, the mother's perception of the child cannot be independent of either her socialization practices or her belief about their outcome. Although the validity of childrearing measures based on interviews is always difficult to assess, the number of associations found due to "halo" effect would probably be lessened when the consequent variable is a systematically quantified measure based on the child's actual behavior as rated by a neutral observer.

Summary

In this study selected childrearing techniques and child behaviors as rated from a mother interview have been related to an experimental measure of resistance to temptation in four-year-old children.

The reported findings and discussion indicate that conformity to rules is a highly complex domain for psychological investigation. Resistance to temptation appears to be related to a number of factors, some of which have been tapped in this study and probably others which will be revealed in the future.

Severity–Permissiveness

There appears to be no simple relation between resistance and the mother's attitude toward severe and permissive training practices. Our results indicate that some measures which seem to reflect heavy socialization demands are more effectual than permissive practices. These measures pertain to activity level of the child during infancy and early childhood, severity of weaning, long toilet training, use of scolding and physical punishment, and infrequent use of "reasoning" for discipline. On the other hand, another set of variables found significantly related to resistance (primarily for boys) appears to indicate a permissive attitude. These practices were explanations about sex, lenient attitude toward cleanliness, and use of withdrawal of love as punishment.

are physically punished more than girls (22, 23). Directly related to resistance to temptation are the results of the dog experiments which indicated that constitutional factors of breed and sex, seemingly related to sensitivity to punishment, were important in determining resistance to temptation (6, 9). Unfortunately, the data from the present study do not permit direct consideration of this hypothesis.

Age and Cognitive Development

An inspection of these two groupings of variables suggested that the first set, which indicates severity, is of greatest relevance during the early period of the child's development. The permissive cluster seems more relevant to a later period in development and seems to require a greater cognitive awareness than would be necessary for the severe training techniques to be effective. These two groupings and the differences between the findings of this study and those of previous reports led to a hypothesis regarding differential child training techniques and cognitive development. This hypothesis is that the more severe, direct, physical forms of childrearing are most efficient during the child's early years and that with increasing cognitive development the more symbolic forms of socialization techniques gain in effectiveness. The generality of this hypothesis is tentative, however, in that the permissive cluster consists of variables related to resistance primarily for boys only. Furthermore, the interpretation regarding differential importance for developmental ages is inferential and has no direct support from this study.

Sex Differences

Readiness to answer questions about sex, lenient attitude toward cleanliness, and use of withdrawal of love were associated significantly with conformity for boys and nonsignificantly with cheating for girls. A home atmosphere of high affectional warmth was related significantly with cheating for girls and nonsignificantly with conformity for boys. Two hypotheses were discussed relevant to these sex differences. Firstly, Whiting's hypothesis that a normally nurturant agent is more effective when administering sanctions is appropriate for boys. Secondly, the association for girls between a "cold" home atmosphere and resistance suggested an identification hypothesis, with girls being more sensitive to their mother's punishment than are boys.

Temptation and Guilt

The associations found by Sears, Maccoby, and Levin (22) have been supported in regard to mother interview measures of the child's evidence of conscience or guilt and of childrearing antecedents. However, the behavioral measure of resistance to temptation tended to be associated with opposite sets of antecedents and to be negatively related to the mother interview measure of conscience. These findings indicate that resistance to temptation and guilt should be studied as distinct phenomena and should not be considered as simply alternative measures of internalized standards.

References

1. Allinsmith, W. The learning of moral standards. Unpublished doctoral dissertation, Univer. of Michigan, 1954.

2. ALLINSMITH, W. The learning of moral standards. In D. R. Miller & G. E. Swanson, *Inner conflict and defense*. Holt, 1960. Pp. 141-176.

3. ALLINSMITH, W., & GREENING, T. C. Guilt over anger as predicted from parental discipline: a study of superego development. *Amer. Psychologist*, 1955, 10, 320. (Abstract)

4. BALDWIN, A. L. *Behavior and development in childhood*. Dryden Press, 1955.

5. BELL, R. Q. Relations between behavior manifestations in the human neonate. *Child Develpm.*, 1960, 31, 463-477.

6. BLACK, A., SOLOMON, R. L., & WHITING, J. W. M. Resistance to temptation as a function of antecedent dependency relationships in puppies. *Amer. Psychologist*, 1954, 9, 579. (Abstract)

7. DINSMOOR, J. A. Punishment: I. The avoidance hypothesis. *Psychol. Rev.*, 1954, 61, 34-46.

8. FAIGIN, H. Child rearing in the Rimrock community with special reference to the development of guilt. Unpublished doctoral dissertation, Radcliffe Coll., 1953.

9. FREEDMAN, D. G. The effects of indulgent and disciplinary rearing in four breeds of dogs. *Science*, 1958, 127, 585-586.

10. GEWITTZ, J. L., & BAER, D. M. The effect of brief social deprivation on behaviors for a social reinforcer. *J. abnorm. soc. Psychol.*, 1958, 56, 49-56.

11. GOETHALS, G. W. A study of the relationship between family esteem patterns and identification, the internalization of values, and aggression of a group of four-year-old children. Unpublished doctoral dissertation, Harvard Univer., 1955.

12. GRINDER, R. E. Behavior in a temptation situation and its relation to certain aspects of socialization. Unpublished doctoral dissertation, Harvard Univer., 1960.

13. GUMP, P. V., & KOUNIN, J. S. Milieu influences in children's concepts of misconduct. *Child Develpm.*, 1961, 32, 711-720.

14. HARTSHORNE, H., & MAY, M. A. *Studies in deceit*. Macmillan, 1928.

15. HARTUP, W. W. Nurturance and nurturance-withdrawal in relation to the dependency behavior of preschool children. Unpublished doctoral dissertation, Harvard Univer., 1955.

16. LIPSITT, L. P., & LEVY, N. Electrotactual threshold in the neonate. *Child Develpm.*, 1959, 30, 547-554.

17. MACCOBY, E. E. The generality of moral behavior. *Amer. Psychologist*, 1959, 14, 358. (Abstract)

18. MACCOBY, E. E. Role-taking in childhood and its consequences for social learning. *Child Develpm.*, 1959, 30, 239-252.

19. MacKinnon, D. W. Violation of prohibitions. In H. W. Murray (Ed.), *Explorations in personality*. Oxford Univer. Press, 1938. Pp. 491-501.

20. Miller, N. E. Learnable drives and reward. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. Wiley, 1951. Pp. 435-472.

21. Schoenfeld, W. N. An experimental approach to anxiety, escape and avoidance behavior. In P. H. Hoch & J. Zubin (Eds.), *Anxiety*. Grune & Stratton, 1950. Pp. 70-99.

22. Sears, R. R., Maccoby, E. E., & Levin, H. *Patterns of child rearing*. Row, Peterson, 1957.

23. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, P. S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-203.

24. Shaplin, J. T. Child training and the identification of pre-adolescent boys with their parents. Unpublished doctoral dissertation. Harvard Univer., 1954.

25. Skinner, B. F. *Science and human behavior*. Macmillan, 1953.

26. Solomon, R. L., & Brush, E. S. Experimentally derived conceptions of anxiety and aversion. In M. R. Jones (Ed.), *Nebraska symposium on motivation*, 1956. Univer. of Nebraska Press, 1956. Pp. 236-248.

27. Whiting, J. W. M. Fourth presentation. In J. M. Tanner & Barbel Inhelder (Eds.), *Discussions on child development: II*. Tavistock Publications, 1954.

28. Whiting, J. W. M. Sorcery, sin and the super-ego: a cross-cultural study of some mechanisms of social control. In M. R. Jones (Ed.), *Nebraska symposium on motivation*, 1959. Univer. of Nebraska Press, 1959. Pp. 174-195.

29. Whiting, J. W. M. Resource mediation and learning by identification. In I. Iscoe & H. W. Stevenson (Eds.), *Personality development in children*. Univer. of Texas Press, 1960. Pp. 112-126.

30. Whiting, J. W. M., & Child, I. L. *Child training and personality: a cross-cultural study*. Yale Univer. Press, 1953.

MILIEU INFLUENCES IN CHILDREN'S CONCEPTS OF MISCONDUCT[1]

PAUL V. GUMP *and* JACOB S. KOUNIN*

Wayne State University

One task of the school is to establish a set of mores or values which will enable it to function effectively. Introduction of this set of school values occurs early in the child's school experience. The investigation to be reported here studied differences in children's ideas of misconduct as these were related to the milieus (school vs. home) in which misconduct could occur. In order to test generality of findings across sex and age variables, boys and girls were interviewed during their first semester in the first grade and again two years later. As a secondary part of the study, the influences of sex and age were also studied.

METHOD

Subjects were 113 boys and 114 girls from six urban schools representing a range of socioeconomic backgrounds. After a brief getting acquainted exchange, children were asked the following questions in the first and, later, in the third grade:

1. What is the worst thing a child could do at school?
2. Why is that so bad?
3. What is the worst thing a child could do at home?
4. Why is that so bad?

---

* College of Education, Wayne State University, Detroit 2, Michigan.

[1] Financial support for this study was received from the College of Education, Wayne State University, and from the National Institute of Mental Health, United States Public Health Service. We are indebted to the personnel of the Detroit Public Schools for making facilities and subjects available. Dr. Bruce Biddle made significant contributions to code development.

A child's answers to these questions can be examined in terms of the kind of misconduct mentioned and the type of explanation given for why the misconduct was bad. Responses of the children were subjected to a comprehensive coding schedule for which a satisfactory reliability had been established; median intercoder agreement for the scales and categories used was 90 per cent. For sake of clarity, codes will be described when results from their use are reported.

Results

Influence of Milieu in Misconduct Type

Misconducts were investigated from three points of view: their general category, their specific type, and their moral seriousness.

Five major categories were sufficient to describe the general categories of misconduct:

1. Violations of rules—no person is directly harmed. "Go visit a friend without asking first." "Talk when you're supposed to study."
2. Assaults on other children. "Pull my sister's hair." "Punch a kid in line."
3. Breaking or damaging property. "Break the kitchen window." "Write on the desks."
4. Nonconformity with adult leader of the setting. "Not mind their mothers." "Don't do what teacher says."
5. Miscellaneous—This includes misconducts which were mentioned infrequently and were too disparate to be combined into one meaningful group. "Steal." "Sneak out." "Tease somebody."

The results of coding children's responses into general categories are displayed in Table 1.

The most consistent difference between school and home on category of misconduct was that related to *breaking*; this category was always most frequent for the home milieu. (When sexes were combined and home and school were compared, p values were .01 for both the first and third grade administrations.)[2] *Nonconformance* tended also to be a home sin but this was most pronounced for the first grade ($p < .01$); the same trend continued for third grade girls ($p < .08$). It is clear that direct disobedience or defiance of teachers was a rarely reported misconduct, but that such affronts to parents were reasonably frequent. A misconduct reported most frequently for school was *assaults* on children. Boys always had more assaults for school than for home ($p < .01$); girls showed this same school-home contrast in the third grade ($p < .02$). The failure of the girls to show more assaults for school than for home in the first grade was related to their heavier use of *violations* for school rather than home misconducts ($p < .001$). This fact, plus other data to be presented, indicated

[2] Differences between settings and between ages are those involving the change of a person's response from one situation to another; for these comparisons the McNemar test was used (2). Differences between sexes were tested with chi square.

TABLE 1

GENERAL CATEGORIES OF MISCONDUCT REPORTED BY CHILDREN

	Violations of Rules	*Assaults on Children*	*Breaking Property*	*Nonconform with Adult*	*Miscellaneous*
First Grade Boys					
Home	38	20	26	11	18
School	46	38	10	2	17
First Grade Girls					
Home	26	22	22	26	18
School	65	21	3	5	20
Third Grade Boys					
Home	50	16	25	3	19
School	41	30	16	4	22
Third Grade Girls					
Home	38	11	27	15	23
School	50	29	12	8	15

that girls' reactions to the impact of school in the first grade differed from boys' reactions.

The qualities of a misconduct are not fully described by the above general category system. Specific misconducts appeared with considerable frequency and seemed highly related to milieu, sex, and age variations. The large use of violations by girls for school derived from frequent reporting of *talking* as a misconduct. Talking was reported by one girl in four in the first grade. By the third grade this had been reduced to one in eight. *Fire play* was another specific misconduct, which showed marked differences between home and school; this was predominantly a boy misconduct, but girls also reported it. Fire play was more frequently a home misconduct for boys and for girls in the first grade ($p < .001$ and .01, respectively) and for boys again in the third grade ($p < .001$). By the third grade one boy in four reported fire misconducts for home—one in 16 for school.

A final aspect of the kind of misconduct related to moral seriousness. The same category may include both morally trivial and morally reprehensible acts. For example, assaults on children can include both: "Sock a kid in the arm and run" and "Beat up little kids who didn't do nothin' to you." There are two aspects of the seriousness question to be reported: extent of *inconvenient* acts and extent of *moral seriousness*.

Inconvenient acts are those which are disturbing to the milieu but which are without direct harm to either the actor or to others and do not violate an important moral taboo. Such acts seem to be natural expressions of chil-

dren's motor and verbal needs. The flavor of this type of sin can best be conveyed by the following examples:

School: "Talk and play." "Run around and make noise." "Throw paper." "Run in the halls."
Home: "Play around at bedtime." "Talk at the table."

In the first grade inconvenient misconducts for school were reported by 42 per cent of the children, home 21 per cent ($p < .001$). In the third grade, inconvenient misconducts were reduced in both settings but the same preponderance of inconvenient sins for school occurred; school 21 per cent, home 11 per cent ($p < .01$).

While inconvenient categorization described a misconduct of a relatively less serious nature, it did not distinguish among the moderately serious and more severely serious misconducts. A scale of general moral seriousness was devised to extend over the whole range of moral seriousness. Results here showed that children in the first grade gave many more truly serious misconducts for home than for school ($p < .01$); by the third grade, however, these serious sins were reported with similar frequency for both home and school.

Influence of Milieu in Misconduct Explanations

A child may explain that a misconduct is bad because (a) it results in some harm to the sinner, (b) it harms other persons, or (c) it is just wrong. (Cases in which the reason given for misconduct being bad was that it was wrong were labeled *reflexive justifications.*) Results from this simple categorization of misconduct explanations are given in Table 2. (Since there were no statistical differences between boys and girls, the data for the sexes were combined.)

Inspection of Table 2 shows that concern for sinner as opposed to concern for others shows no influence of milieu. The only consistent distinction between home and school was that relating to reflexive justification; expla-

TABLE 2

CHILDREN'S EXPLANATIONS FOR WHY A SIN IS BAD

	Harm to Sinner	*Harm to Other*	*Reflex. Just.*
First Grade			
Home	120	81	26
School	110	73	44
Third Grade			
Home	95	113	19
School	82	113	32

nations of "it's bad because it's bad" were most frequently used for school misconducts ($p < .01$ and .05). Inspection of the data showed that the higher proportion of reflexive justifications at school was associated with their relatively greater use in connection with rule violations. In the first grade, for example, reflexive justifications were attached to 24 of the school rule violations, but to only four of the home violations. Third grade data showed the same trend.

It was possible to categorize reasons for misconduct badness more specifically. The frequently used explanations fell into the following categories:

Adults suffer. "My mother would have to clean up the mess."

Children suffer. "Kid could be hurt."

Adults punish. "I'd have to stay inside for a week."

Reality punishes—No person is made responsible for the punishment; the harm to the sinner is a natural outcome of his behavior. "If you don't look both ways you could get hurt real bad." "You could get behind in your work."

Reflexive Justifications and other.

(Sex differences were statistically insignificant, and the sexes were combined.)

Two milieu contrasts were consistent in Table 3. First, only home adults suffered as a result of a child's misconducts; teachers almost never suffered. Secondly, children at school suffered more frequently than children at home ($p < .001$). Apparently adults punished with similar frequency in both home and school. In the first grade, reality punishments are much more frequent for home than for school ($p < .001$). Reality punishments at home probably reflected the large number of dangerous situations at home about which younger children are constantly warned: playing in street, climbing tress, etc.

Just as milieu influenced whether or not the central adult suffered, it also influenced the type of punishment this central adult administered. In

TABLE 3

CHILDREN'S SPECIFIC EXPLANATIONS FOR WHY A SIN IS BAD

	Adults Suffer	*Children Suffer*	*Adults Punish*	*Reality Punishes*	*R. J. and others*
First Grade Children					
Home	42	37	50	58	40
School	10	59	54	26	78
Third Grade Children					
Home	60	46	28	64	29
School	1	92	16	54	64

the first grade, parents were reported as using corporal punishments in 64 per cent of their punishments, teachers 22 per cent ($p < .001$). By the third grade punishment by adults was markedly reduced, but home tended to maintain the greater proportion of corporal punishment: parents 32 per cent, teachers 19 per cent ($p < .10$).

Sex and Age Influences in Misconduct Type

For purposes of clarity, differences in misconduct ideas relating to sex or age have been so far ignored. However, certain findings indicate that both age and sex sometimes operated in relation to milieu influences and that age was occasionally relatively independent of milieu.

The conjunction of milieu, sex, and age variables is particularly striking in connection with the categories of misconduct in Table 1. In the first grade, boys' and girls' responses in Table 1 were quite similar for home ($p > .20$). First grade school, however, created sharp differences between boys and girls ($p < .01$). Here girls emphasized the school rules (often talking), and boys were concerned with assaults and with breaking. In the first grade, girls accepted the emphasis the school puts on rule violations; at this time and place the more aggressive concerns of assaults and breakings were diminished. However, by the third grade, the pattern of school misconduct for boys and girls was roughly similar ($p > .40$).

A second boy-and-girl difference relating to milieu was the girls' heightened concern with nonconformance to parents. In the first grade, girls had 26 home nonconformances to boys' 11; in the third grade, girls had 15 to the boys' 3 ($p < .01$ in both cases).

Boy-and-girl differences in *fire play* were impressive only for the home; although boys showed a weak tendency to have more of these at school than girls, it was only in the home responses that boys' *fire play* shows a decidedly greater frequency (p's $< .01$). Finally, as noted before, talking in school tended to be a girls' problem ($p < .01$).

The only impressive change in misconduct type related to age has to do with aspects of seriousness. As might be expected, inconvenient misconducts decreased with increasing age ($p < .01$), and general moral seriousness increased with age ($p < .01$); these results held for both home and school.

Sex and Age Influences in Misconduct Explanations

No consistent differences were found between boys' and girls' explanations as to why a misconduct was bad.

Differences in age did relate to misconduct explanations. It is clear from Table 2 that as the child gets older he increases his concern with harm to others. This change is consistent over settings: concern for others at home rises from 32 to 50 per cent, at school from 36 to 50 per cent (p's $< .01$).

If Table 3 is used to compare age changes, it can be shown that the increased concern about others at home resulted from an increase in adult

suffering ($p < .05$), and the increased concern for others at school derived from an increase in child suffering ($p < .001$). Table 3 also shows that adult punishment as a reason for avoiding a misconduct is markedly reduced as children mature. Again the difference is consistent across settings: adult punishment at home drops from 22 to 12 per cent; at school, from 24 to 7 per cent (p's $< .01$). The type of punishment expected from adults also changes with age: corporal punishment for home drops from 32 to 9 per cent; for school, from 12 to 3 per cent ($p < .001$ and .02).

Since the data here presented are somewhat complicated, a listing is offered to summarize the major relations of milieu, sex, and age variables to the indices employed.

Milieu Contrasts:

School misconducts emphasized assaults on children and inconvenient and talking misbehaviors.

Home misconducts were weighted toward breaking of objects, fire play, and nonconformance with parents.

School explanations emphasized reflexive justification, children suffering, and noncorporal punishment.

Home explanations emphasized adult suffering and corporal punishment.

Sex Contrasts:

Girls emphasized talking at school; boys, fire play at home.

Girls also emphasized noncomformance with parents.

Girls in the first grade were dominated by concern with school rule violations while boys were more likely to give assaults and breakings; this was a temporary reaction which disappeared by the third grade.

Boys and girls did not differ in the type of explanation they offered for misconducts.

Age Contrasts:

As children became older, they increased their offering of morally serious misconducts and decreased stating inconvenient ones.

With age development, children's explanations referred more frequently to harm to others and less frequently to adult punishment; adult punishment, when expected, contained less corporal punishment.

Discussion

Children probably answer the question, "What's the worst thing a child can do?" with a report of their preoccupations. It is not likely that our subject's answers would have been the same if they were presented with alternatives of different moral seriousness. Given a choice, most children would probably say that "stabbing someone" was worse than "talking while the teacher is out of the room."

If the misconduct ideas are taken to represent preoccupations, it should be possible to infer from these preoccupations the impact of the school or home milieu upon our subjects. Their preoccupations should reflect (a) the

needs which they want to express but which the milieus find alien, (b) the temptations or opportunities which the milieu offers, and (c) the expectations held about what gets one into trouble. Preoccupations should change over time as some misconduct issues become settled and other issues take their place.

Using such an inferential approach with regard to milieu contrasts, it might be assumed that school is perceived as a place in which one's natural inclinations toward such things as motion and noise (inconvenient misconducts) are most heavily suppressed. The problem of handling one's aggressive impulses toward other children is also seen as particularly acute at school. There is probably a double reason for this: first, many other children are ever present at school, presenting a constant temptation—or opportunity; second, school will not tolerate the aggressive interactions, many of which are either excused or unnoticed at home.

Home is perceived as a place where a freer range of behavior is possible and where certain adventurous sins can result. At home a child has more occasion to be without close adult supervision; he can avail himself of more adult props and use them more freely. From this relative freedom come more misconducts dealing with fire play and breaking. One may also infer that parents and teachers occupy quite different roles in the child's mind. While teachers seem to be relatively unassailable (few nonconformances), parents seemed more vulnerable (more nonconformances).

Milieu influences also extend to explanations about why a misconduct is bad. There are indications that children have difficulty in seeing any good reason for some of the school rules. This deduction was supported by the fact that children reported more reflexive justifications for school, and these reflexive justifications were especially associated with school rule violations. Apparently, at school, one simply takes some prohibitions on faith.

The relatively unassailable role of teachers seems also related to misconduct reasons. Children rarely perceived school adults as suffering harm from their misconducts; they perceived home adults suffering such harm very frequently. Although the parent is seen as more vulnerable, the data on corporal punishment would indicate that he or she was also seen as inflicting more primitive, and probably more painful, punishment. One is not inclined to psychologize over this issue: parents most probably use corporal punishment more frequently than teachers.

One might hope that the school, as a socializing agent, would develop an increasing concern for others; this would be reflected in the present data by school explanations which showed more interest in harm to others and less concern about adult punishment. As reported, however, these improvements in socialization were a function of age and operated for both settings. What the school did do was to determine the kind of others with which the child became concerned: school increased children's concern with what happens to other children while the home increased concern with what happens to other adults. One might argue that the very large increase in child

suffering at school represents a socialization gain superior to the rise in adult suffering at home. The reasoning here would be that a broadened concern about others of equal power (children) represents a more justice oriented outlook than increased interest in others of superior power (parents).

The results relating to sex differences indicated that these were few, specific, and related to misconduct type—not to misconduct explanations. One might have expected some impressive sex differences because it is commonly accepted that boys have more trouble adjusting to school than girls. In one large city, referrals to visiting teachers of boys with school conflict problems outnumbered referrals of girls by a ratio of five to two (1). One might expect school conflict among boys to result in boys offering a pattern of misconduct concerns different from the one offered by girls. This expectation was born out in the first grade school data: girls seemed to take on the school emphasis of obeying rules; boys became more concerned with assaults. However, by third grade, boys and girls showed no statistically significant difference in misconduct categories; girls reduced their concern with violations of rules and increased their assaults—they become like the boys. Whatever marks the difference between boys' and girls' adjusment in school, it seems unrelated to ideas about misconduct in school. Future research might investigate the question of whether this same sex similarity continues in the higher grades.

The constant sex differences of girls' concern with nonconformance at home and with talking at school and boys' interest in fire play at home are reported as matters of possible interest without attempts at explanation.

Age changes would indicate that, over time, children show their social maturation in two ways. First, they become able to see beyond the inconvenient sins that often get them into trouble to misconducts of more moral significance. Secondly, they begin to drop concern with obvious self-centered consequences, like adult punishment, and become more interested in what happens to other persons.

Summary

Children in the first and, later, in the third grades were asked which misconducts they considered the worst at school and at home; they were also asked to explain why they considered these misconducts bad. Data were analyzed in terms of how relevant milieu was related to the qualities of the misconducts and to their explanations. Comparisons were made across variables of sex and age. Results showed that relevant milieu related both to the kinds of misconducts reported and to explanations for misconducts. Some of these relations were consistent regardless of the age or sex of the child; others obtained only for specific sex or age groups.

Differences in misconduct ideas related to sex were confined to kinds of misconduct—not to explanations—and were limited to those occurring in

one milieu or the other. Differences related to age tended to hold for both sexes and in both milieus.

Hypotheses were offered concerning the ecological and maturational variables responsible for the obtained relations.

References

1. Baker, H. J. *Visiting teacher program.* Psychological Clinic, Detroit Public Schools, 1958-59.
2. Siegel, S. *Nonparametric statistics for the behavioral sciences.* McGraw-Hill, 1956.

MEASUREMENT OF CHILDREN'S SENSITIVITY TO AUDIENCES[1,2]

ALLAN PAIVIO,* ALFRED L. BALDWIN, *and* SEYMOUR M. BERGER[3]

Cornell University

In our culture at least, social rewards and punishments are generally contingent on public exposure. As Heider (4) points out, what is perceived is often judged favorably or unfavorably. It is not surprising, therefore, that a public performance may involve elements of either attraction or fear, or both, for the "performer." Thus, we observe that some individuals "seek the limelight," others react with shyness, embarrassment or "stage fright" in the presence of observers and characteristically avoid public situations. Others apparently manifest both reaction tendencies to a high degree, e.g., professional actors who report extreme stage fright despite years of stage experience. Hollingworth (5) emphasized this conflictful nature of public exposure, referring to the "striking paradox [of] the joint activity of two opposing impulses—craving for an audience and fear of it" (5, p. 183). The "paradox" is significant if not surprising. Arousal of one or both of the opposing tendencies may be reflected in the quality of an individual's public performance as well as in his willingness to appear at all before an audience, and adequate measurement of the strength of these tendencies should enable us to predict individual differences in audience effects.

The present report describes the development and initial validation of a questionnaire designed to measure these conflicting tendencies in children. This questionnaire was initially based on an adult inventory developed by Paivio and Lambert (10) for the measurement of audience sensitivity, a concept defined in their study as a predisposition to be anxious in audience situations. However, in addition to describing anxiety and avoidance of public exposure, the children's questionnaire attempts to measure the motive to seek an audience, thus introducing the possibility of an approach-avoidance conflict analysis of behavior in audience situations. Such an analysis was employed in a separately-reported study (8), where children classified

---

* Department of Psychology and Sociology, University of New Brunswick, Fredericton, N. B., Canada.

[1] This research was supported by a grant (M-901) from the National Institute of Mental Health.

[2] We wish to thank Dr. Mary Gallwey and Irene Stein for their help with aspects of the study.

[3] Now at Indiana University.

Child Develpm., 1961, 32, 721-730.

(on the basis of scores on the children's questionnaire) as having strong tendencies both to seek and avoid public exposure made more speech "errors" before an audience than did less conflicted subjects. Further theoretical considerations and related research are more fully discussed elsewhere (7, 9, 10).

Development of the Questionnaire

Initial Form of the CASI

The adult Audience Sensitivity Inventory (10) consists of 21 items referring to typical reactions to various audience situations. Items from this scale were reworded to be more suitable for children and some new items added, forming a 31-item Children's Audience Sensitivity Inventory (CASI).

The CASI was mimeographed and group-administered to 111 fourth-, fifth-, and sixth-grade children from an urban elementary school. Procedural instructions and the items were read aloud by an *E*, and the children responded by checking one of two answer spaces (labeled "Yes—like me" and "No—not like me") following each item.

Twenty of the items showed enough variation to be subjected to further analysis. A matrix of phi coefficients, computed between the dichotomously-scored items, was factor analyzed to simple structure, following Thurstone's (13) centroid method. The analysis yielded two clearly orthogonal factors. The first showed high loadings on seven items referring to public exposure of the self or a self-product (e.g., I like to tell a story in front of the class, I like to show my work to my classmates, etc.) and was labeled Exhibitionism (Exh). The second was defined by six items referring to negative attitudes toward public exposure (e.g., I feel bad when someone sees me goof) and was called Self-Consciousness (S-C).

Assuming that the Exh factor measures an audience-seeking tendency and S-C apprehension with respect to public exposure, orthogonality of the factors meant that *S*s scoring high on both (i.e., *S*s for whom a public performance presumably involves an approach-avoidance conflict) could readily be selected for research purposes. However, examination of the two factors suggested caution in their interpretation. The items comprising Exh are relatively specific in their referents, whereas the S-C items are more global and less situation-specific, suggesting that the factors may not be of comparable validity as predictors of reactions to particular audience situations. Some evidence that this might be the case was obtained in a separate study (9) concerned with the antecedents of audience sensitivity. One hundred and eighty-seven third and fourth grade children wrote compositions on either "Why I like to recite in front of the class" or "Why I do not like to recite in front of the class," the children being free to choose either theme. Liking reciting, as expressed by the composition theme, correlated significantly (biserial $r = .67$, $p < .0001$) with Exh, but only negligibly with S-C. In itself this finding suggested only that the composition theme could be

interpreted as a further Exhibitionism "item." However, analysis of the content of the compositions revealed that 85 per cent of the compositions written on "Why I do not like to recite" included statements clearly referring to anxiety as the reason for not liking reciting (e.g., "because I shake all over," "because I am scared," etc.). Thus, while Exh correlated negatively with direct statements of anxiety concerning recitation, S-C was apparently unrelated to such references and could not be unequivocally interpreted as measuring fear-mediated avoidance of public exposure.

The above considerations led to a revision of the CASI to include a third scale designed to measure situationally-specific avoidance tendencies regarding public performance.

Revised CASI

Items. Sixteen items which refer explicitly to anxiety reactions in, or avoidance of, the kinds of situations represented by the Exh items were added to the CASI. Some of these were direct modifications of Exh items to include a fear reference; others were reproductions of popular statements made by *Ss* who wrote compositions on "Why I do not like to recite" in the study (9) cited above. The new "scale" was labeled Audience Anxiety (AA).

In the course of the revision, six new items were added to the Exh factor to increase its reliability. The revised CASI thus consists of 13 Exhibitionism, 6 Self-Consciousness, and 16 Audience Anxiety items. The items are listed below.

Exhibitionism items:

I like to recite poems in front of other people.
Everytime I get a chance to do something in front of the class I take it.
In school, I always raise my hand if I know the answer.
Even if I know the answer, I usually do not raise my hand. (—)[4]
If I wrote a prize-winning poem I would rather have someone else read it in front of the class. (—)
If my paper is hung on the bulletin board, I'd rather *not* have my name on it. (—)
I like to show things I make to other children.
I would like to make something while the whole class watches.
I like to show my work to my classmates.
I like to sing in front of others.
I like to tell a story in front of the class.
I would like to be on the stage in front of many people.
If my paper is hung on the bulletin board, I like everyone who sees it to know it's mine.

Self-Consciousness items:

I am bashful with most strangers.
I feel funny when I walk into a room full of people.
I often wonder what others think of me.
I feel bad when someone sees me goof.

[4] Items followed by a minus sign (—) are scored if the response is negative.

I often worry about what people think of me.
Other people can hurt my feelings easily.

Audience Anxiety items:

I worry that I might forget my lines when I recite poems in front of the class.
I do not like acting in plays because I am afraid that people will laugh at me.
I feel nervous when others look at work I have done.
I feel embarrassed when I am asked to sing in front of others.
My knees shake when I recite in class.
I worry that I will make a mistake when I give a report in front of the class.
I never feel one bit nervous when I recite. (—)
When it is my turn to get up and recite in class, I feel my heart pounding hard.
I do not like reciting in class because I might make a mistake and others would laugh at me.
I am afraid that other children will laugh at me when I show my work to them.
Sometimes I shake all over when I am asked to recite in class.
Sometimes my voice shakes when I recite in class.
Once I had to do something in front of an audience and I was really scared.
I feel scared when I recite in front of the class.
I am very nervous if the whole class watches me while I am making something.
When I recite in class I often wonder what others are thinking about me.

Reliability of the revised CASI. Test-retest (six-month interval) and split-half reliabilities of the Exh, S-C and AA scales were computed for a sample of third-, fourth-, and fifth-grade children from two elementary schools. The questionnaires were group-administered by teachers, following detailed instructions provided by an *E*. The reliabilities are presented in Table 1.

TABLE 1

TEST-RETEST AND SPLIT-HALF RELIABILITIES OF THREE SCALES (THIRD-, FOURTH-, AND FIFTH-GRADE CHILDREN)

	N	*Exh*	*S-C*	*AA*
Test-retest *r**	209	.62	.45	.54
Split-half *r*†	223	.67	.66	.80

* Six-month interval.
† Corrected by Spearman-Brown formula.

The test-retest correlations are low but not necessarily unsatisfactory, considering the long retest interval. They may in part reflect intervening classroom experiences which affected the motivational predispositions the

scales were intended to measure. The problem requires separate investigation, however, and extensive speculation is unnecessary here.

Correlations between CASI, Manifest Anxiety, and Test Anxiety Scales. While mainly concerned with the relations among the three CASI scales, we were also interested in comparisons with instruments on which published data were available. Thus, the items of the children's forms of the Manifest Anxiety (CMAS) (1) and Test Anxiety (TA) (11) Scales were added to the CASI. The composite questionnaire was administered to three samples of children from school grades 3 through 10 (as part of separate studies, the purposes of which are not relevant here), including the sample on which the above reliability data are based. Intercorrelations of the five scales were computed, separately for boys and girls, for each of the samples. In some instances there is considerable variability in the individual sample correlations between pairs of scales, but the over-all pattern of relations among the five scales is similar in the different samples and their data will not be individually presented here. In order to give a representative picture of the interrelations, the data were pooled (total $N = 421$) using Fisher's z'-transformation method (*see* 3) for computing average correlations. These correlations are presented in Table 2.

TABLE 2

AVERAGE INTERCORRELATIONS AMONG FIVE QUESTIONNAIRE SCALES

	Exh	*S-C*	*AA*	*TA*	*CMAS*
Exh		—.21	—.55	—.40	—.27
S-C			.58	.59	.61
AA				.74	.62
TA					.70

NOTE.—Averages are based on six subgroup *r*s (three samples, boys and girls treated separately) and computed by Fisher's z'-transformation method (total $N = 421$). Homogeneity of subgroup *r*s not demonstrated in every case. All average *r*s are significant, $p < .01$.

Considering, first, the relations among the CASI scales, Table 2 shows that Exh and S-C are only slightly negatively correlated, as in the factor analysis of the initial CASI. However, audience anxiety is apparently a component of both Exh and S-C, for AA correlates moderately with each (i.e., negatively with Exh, positively with S-C). To the extent that such inferences are warranted from questionnaire data, the Exh-AA correlation suggests that attraction and fear vis-à-vis public performance are neither bipolar nor independent. A subject could score moderately highly on both scales, presumably indicating that he is somewhat conflicted in his attitude toward audiences, but he is unlikely to be at once extremely exhibitionistic and highly fearful of public exposure. Thus, considering at this point only the

intercorrelations of the CASI scales, an approach-avoidance conflict analysis of audience-oriented behavior would be possible using either AA or S-C scores as measures of a dispositional avoidance tendency, but would clearly be more feasible with the latter in view of its relative independence of Exh.

The correlations among the anxiety scales and S-C are all moderately high. Most notable among these (particularly since the two scales were independently developed for somewhat different purposes) is the correlation between AA and TA, which consistently ranges between .70 and .80 for the different subsamples (average $r = .74$). This and other aspects of the questionnaire data will be further discussed at a later point in the paper.

Behavioral Validation of the Revised CASI

The following study, undertaken to test the validity of the CASI scales, takes advantage of a natural setting in which children were asked to volunteer for a public performance. It was assumed that, if the Exh factor is a valid measure of an audience-seeking tendency, high scorers on Exh should volunteer more readily for a public performance than *Ss* scoring low on Exh. Similarly, to the degree that the S-C and AA scales measure the tendency to avoid public exposure, scores on these scales will correlate negatively with "willingness to volunteer."

Subjects

The *Ss* were 56 boys and 54 girls attending (at different times) a summer camp near Ithaca, N.Y. The boys ranged from 10 to 16, the girls from 8 to 15, years of age.

Procedure

Boys. The boys' group consisted of five camp units, the number of boys in a unit ranging from 8 to 18. Their camp program included a "skit" night for which each unit was to prepare a public performance and for which the campers were asked by their unit leaders to volunteer. Participation was optional, but considerable social pressure was presumably involved. The unit leaders were asked to rate each boy in their respective units on eagerness to volunteer for the skit, using scoring sheets containing the following instructions:

> Please list below the names of those boys who volunteered to perform in a skit, and indicate, to the best of your knowledge, how *readily* each boy volunteered by writing the appropriate number (1, 2, 3, or 4) in the column headed "Willingness to Volunteer," where
>
> 1 = volunteered readily (appeared eager to participate)
> 2 = volunteered (but not particularly eagerly)
> 3 = volunteered reluctantly (needed some coaxing)
> 4 = volunteered *very* reluctantly (group pressure required)

Not volunteering at all represented a fifth category.

The CASI, CMAS, and TA scales were completed by each boy on the day of the skits. They were administered by the camp counselors who followed mimeographed instructions concerning the procedure.

Girls. A procedure similar to the boys' was followed with the girls' sample (four camp units of 10 to 18 *Ss*), but more carefully controlled. The skit night was included in their camp program at the request of an *E* who was able to meet with the girls' counselors and discuss the study in detail, and thereby establish uniformity in the procedure. Criteria were suggested for rating the children: *Ss* who volunteered immediately after the first call for volunteers were to be scored 1 ("Volunteered Readily"), those who agreed to participate after a second and a third request were scored 2 and 3 respectively, and those who volunteered only after the greatest insistence by the counselor and other campers were scored 4 ("Volunteered very reluctantly").

The questionnaires were completed by the girls on the day following the skits.

Results

It would have been desirable to consider each camp unit individually, but this was impractical with the boys' data. In view of the small numbers in some of their groups and the fact that two of the counselors used only a two-point scale in rating the boys on volunteering (1 and 2 in one case, 1 and 5 in the other), the boys' volunteering scores were dichotomized at a point where a common division was possible and the units combined. The resulting two groups to be compared consisted of 27 *Ss* rated as "volunteering readily" and 29 rated as volunteering less readily or not at all. Biserial correlations were computed between these dichotomized volunteering scores and each of the CASI scales.

While the procedure for the girls' study differed somewhat from that of the boys, the conditions for the two studies seemed sufficiently similar to make comparisons meaningful. Thus, the girls' data were analyzed in the same manner as those of the boys. The resulting biserial correlations between willingness to volunteer for the skits and CASI, TA, and CMAS scores, for both sexes, are shown in Table 3.[5]

All correlations are in the expected direction, i.e., volunteering for public performance correlates positively with Exh and negatively with S-C and AA (as well as with CMAS and TA).[6] However, conventional significance

[5] Since all but one of the girls' counselors used the four categories of the scale in rating the *Ss* in their units on volunteering, an alternative analysis was feasible with the girls' data. Product-moment correlations between the questionnaire scores and volunteering ratings were computed for each of the groups and combined into single correlation estimates, after establishing homogeneity of the individual *r*s, following Fisher's (3) technique. The resulting average correlations are generally comparable to the girls' biserial correlations shown in Table 3 and will not be separately discussed.

[6] In view of their intercorrelations (Table 2) the scales are obviously not independent predictors of volunteering.

TABLE 3

BISERIAL CORRELATIONS BETWEEN RATED "WILLINGNESS TO VOLUNTEER FOR PUBLIC PERFORMANCE" AND SCORES ON FIVE QUESTIONNAIRE SCALES

Sex	*N*	*Exh*	*S-C*	*AA*	*TA*	*CMAS*
Boys	56	.31*	—.23	—.35*	—.36*	—.29*
Girls	54	.13	—.20	—.45**	—.39**	—.26

* $p < .05$, one-tailed test.
** $p < .01$, one-tailed test.

levels are not reached in every case. The AA and TA correlations with volunteering are significant ($p < .025$, one-tailed test) for both boys and girls, Exh and CMAS for boys only ($p < .05$), while the S-C correlation is not significant in either sample. It should be noted, however, that none of the sex differences is significant, increasing confidence that correlations in the predicted direction obtain.

Particularly noteworthy is the correlation of —.45 ($p < .01$) between girls' AA scores and volunteering, which indicates that at least 20 per cent of the variance in rated volunteering scores is accounted for by "audience anxiety."

DISCUSSION

The results of the preliminary validating study generally support the validity of the Children's Audience Sensitivity Inventory. The accuracy of prediction leaves much to be desired from a practical point of view, but the Audience Anxiety scale correlations with volunteering (particularly in the girls' sample) compare favorably with validity coefficients of personality tests in general (*see*, e.g., 6). The Exhibitionism scale was less successful as a predictor, only the boys' correlations with volunteering being significant. Possibly the Exh items are less reliable predictors for girls because skill in the activities they imply (singing, reciting, acting, etc.) is more valued by girls (cf. 12) and, hence, social acceptability plays a more important role in determining girls' than boys' responses to the items. There may also have been some effect from the fact that the CASI was completed by the boys before, and by the girls after, actual participation in the skit; i.e., experience in the skit may have affected the girls' responses to the Exhibitionism items but could not have done so in the case of boys. However, as already pointed out, the sex differences in the correlation between Exhibitionism and volunteering is not significant, and extensive speculation is unnecessary.

The Self-Consciousness scale also did not correlate as highly with volunteering as did Audience Anxiety. The difference may be due to the smaller number of items in the S-C scale and/or to the specific relevance of the AA

items to the criterion situation in the present study. In any case, it does not justify assuming lower validity of the S-C items with respect to other kinds of audience situations.

The relation between Audience Anxiety and Test Anxiety is of some interest here. The high correlation between them and the fact that, of the scales included in the questionnaire, they correlate most highly (and approximately equally) with volunteering, suggests common anxiety-arousing features in the situations implied by the two sets of items. The crucial element may simply be the classroom situation, to which most items of both scales make reference, but the fact that the criterion situation in the present study was *not* a classroom suggests something beyond this.

Both the test and the audience situation involve, implicitly or explicitly, evaluation by others. Test Anxiety and Audience Anxiety scores may therefore reflect a generalized "sensitivity" to social evaluation founded on prior experience in which behavior has been devalued and punished. Such an interpretation is supported by the findings of a study on antecedents of audience sensitivity (9): children whose social behavior and achievements were evaluated relatively unfavorably by parents, and who were punished frequently for failure to meet parental standards, had significantly higher AA (and TA) scores than children who were favorably evaluated and infrequently punished. The same interpretation may hold for the Manifest Anxiety Scale, which contains a number of items having reference to social evaluation (e.g., "I feel that others do not like the way I do things"). However, while moderately intercorrelated, indicating some general factor, an item analysis (not reported here) shows that the items of each scale correlate more highly with their "own" scale than with the others, suggesting situational specificity in anxiety (cf. a study by Dixon *et al.* [2], in which a factor analysis of a "social anxiety" scale yielded a strong general factor and four group factors). The basis of such specificity cannot be determined from existing data.

The concepts employed in this study raise many questions, both regarding their interpretation and their relation to such concepts as achievement need, self-esteem, etc. Such questions and specific lines research might take to answer them are discussed elsewhere at length (7, 9) and shall not be introduced here. Ultimately, the value of these concepts depends on how far they will take us in the description and conceptualization of social influence, particularly that aspect of social influence which is attributable to observing, evaluating, and potentially approving or disapproving others. The instruments described above have already proven useful in that regard (*see* 8, 9), and further research is currently under way.

Summary

A Children's Audience Sensitivity Inventory was developed to measure individual differences in motivational reactions to the presence of an audi-

ence. The Inventory includes three scales labeled Exhibitionism, Self-Consciousness, and Audience Anxiety.

Empirical validation of the inventory involved a camp situation in which children volunteered for performance in a skit. It was predicted that Exhibitionism would correlate positively, and Self-Consciousness and Audience Anxiety negatively, with willingness to volunteer (as rated by camp counselors). The findings generally supported the predictions, most strongly in the case of Audience Anxiety. Relations to the children's Test Anxiety and Manifest Anxiety scores are also discussed.

References

1. Castaneda, A., McCandless, B. R., & Palermo, D. S. The children's form of the manifest anxiety scale. *Child Develpm.*, 1956, 27, 317-326.
2. Dixon, J. J., De Monchaux, C., & Sandler, J. Patterns of anxiety: an analysis of social anxieties. *Brit. J. med. Psychol.*, 1957, 30, 107-112.
3. Edwards, A. L. *Experimental design in psychological research.* Rinehart, 1956.
4. Heider, F. *The psychology of interpersonal relations.* Wiley, 1958.
5. Hollingworth, H. L. *The psychology of the audience.* American Book, 1935.
6. Kimble, G. *Principles of general psychology.* Ronald, 1956.
7. Levin, H., & Baldwin, A. L. Pride and shame in children. In M. R. Jones (Ed.), *Nebraska symposium on motivation.* Univer. of Nebraska Press, 1959. Pp. 138-173.
8. Levin, H., Baldwin, A. L., Gallwey, M., & Paivio, A. Audience stress, personality and speech. *J. abnorm. soc. Psychol.*, 1960, 61, 469-473.
9. Paivio, A. Child rearing antecendents of audience sensitivity. Unpublished doctoral dissertation, McGill Univer., 1959.
10. Paivio, A., & Lambert, W. E. Measures and correlates of audience anxiety ("stage fright"). *J. Pers.*, 1959, 27, 1-17.
11. Sarason, S. B., Davidson, K., Lighthall, F., & Waite, R. A test anxiety scale for children. *Child Develpm.*, 1958, 29, 105-113.
12. Terman, L. M. Psychological sex differences. In L. Carmichael (Ed.), *Manual of child psychology.* Wiley, 1946. Pp. 954-993.
13. Thurstone, L. L. *Multiple factor analysis.* Univer. of Chicago Press, 1947.

PERSONALITY AND BEHAVIOR PATTERNS OF CHILDREN WITH FUNCTIONAL DEFECTS OF ARTICULATION[1]

ARTHUR L. SOLOMON*

Chico State College

It is reported that the outstanding majority of speech defects in children are articulatory in symptom and functional in origin (11, 17), although no clear-cut etiological picture has as yet been established to account satisfactorily for the occurrence of the defects or for their subsequent disappearance. Studies of somatic variables over the past 20 years have not been especially rewarding (13), while clinical evidence has increasingly stressed the emotional and personality factors found as concomitant symptomatology in children with various functional speech disorders described as "baby talk," "infantile perseveration," "delayed speech," "speech inhibition," or "mutism" (1, 3, 4, 6, 12, 15).

Experimental studies using personality inventories have been either inconclusive or contradictory, although the investigations of Wood (16), Nelson (10), and FitzSimons (5) of children with functional defects of articulation, using rating scales, controlled observations, or projective techniques, have yielded more conclusive findings. Beasely (2), McCarthy (7), and Mowrer (9), particularly, have emphasized the importance of the parent-child rela-

* Speech and Hearing Clinic, Chico State College, Chico, California.

[1] Adapted from a doctoral dissertation, Stanford University, 1960. The author wishes to acknowledge the generous assistance of Dr. Virgil A. Anderson, Director of the Division of Speech Pathology and Audiology of the Stanford School of Medicine; Dr. Dorothy Huntington; and Dr. Lois Meek Stolz, Professor of Psychology Emeritus, Stanford University.

Child Develpm., 1961, 32, 731-737.

tionship in the acquisition and adequacy of language, as well as other social learnings.

It would appear to follow that disturbances in the socialization process due to certain unfavorable environmental factors and showing up as problem behavior or unfavorable personality traits would, because of the sensitivity of the speech mechanism to disturbances in interpersonal relationships, be reflected in communicative disturbances as well. This hypothesis is apparently supported by the observation that speech defects of a functional nature follow patterns very similar to those of the usual childhood behavior problems which reach their ascendancy in the preschool years and gradually diminish or become displaced during the early years of school (8, 14). If these implications are correct, behavior problems, or specific personality traits and their relation with speech disturbance, should then be able to be experimentally demonstrated. If demonstrated, some further evidence will be added to the supposition that functional speech defects in children are part of a total adjustive pattern and not disparate from other social learnings.

A study was consequently designed to explore the behavior and personality traits of children with functional defects of articulation to determine whether children with these defects differed from their normal-speaking peers in specific behavior problems; whether the severity and degree of that difference could be determined; whether the behavior and personality traits of these children followed any particular pattern; and whether there might exist a correlation between the severity of the articulatory defects and the severity of the behavior disturbances, if any.

Method

Forty-nine boys and girls, of a comparatively homogeneous socioeconomic background, with functional defects of articulation, were matched in pairs with a control group of normal-speaking children on the basis of age, grade, sex, and intelligence. The experimental group represented a broad sampling of the speech defective population as no selection was made on the basis of severity of speech involvement. For inclusion in the experimental group, the subjects had to exhibit at least one consistently defective sound; they had to have scored 90 or above on the nonverbal scale of the California Short-Form Test of Mental Maturity: they had to be in the first grade; not a member of a twin pair; they had to have good health, good hearing, and a negative medical history; they had to have normal speech structures; and a mother who was willing and able to participate in the study. For inclusion in the control group, the subjects had to have normal articulatory proficiency and had to be within 90 calendar days and within five points of their experimental counterparts in age and IQ, respectively, and of the same sex.

In order to ascertain the existence and degree of problem behavior and to evaluate personality differences, mothers of both groups were interviewed

and asked to describe their children's behavior, over the past two years, in nine representative areas of child development and personality. These areas consisted of (a) eating behavior, (b) sleeping behavior, (c) toilet training, (d) fears and anxieties, (e) comfort patterns, (f) tension, (g) aggression, (h) dependency, and (i) peer relations. A parent interview-questionnaire was devised to elicit information in these areas and, in general, the interviews followed similar patterns: (a) a general question to encourage the parent to describe her child's behavior in a given category and (b) specific questioning to determine frequency or severity of the reported behavior.

The interviews were recorded on tape and later typed verbatim from the recordings and so arranged that each of the nine behavior areas could be detached and rated separately. The protocols were rendered anonymous, and experimental and control data were randomized so that two judges, working independently, could rate all the children on each specific behavior area and again on each child's total behavior to determine an "over-all adjustment score." The ratings of the two judges, using a five-point scale, were averaged, yielding a combined score. Coefficients of agreement between the ratings of the judges were calculated along with a test of the consistency and range of their respective ratings. The investigator was one of the two judges participating in each rating procedure.

To determine the relationship between articulation and adjustment, the experimental subjects were divided by the investigator into two categories on the basis of their articulatory errors: (a) those with mild-moderate speech involvement and (b) those with marked-severe involvement. Then, these two groups were compared on their over-all adjustment scores. To determine characteristic behavior patterns, four categories were decided upon that might yield significant data, namely: (a) orality, (b) passivity-submissiveness, (c) internalizing behavior, and (d) externalizing behavior. The items under "orality" included thumb or finger sucking, nail biting, excessive crying, chewing objects, etc. The second category of "passivity-submissiveness" contained such examples of behavior as the following: tries hard to please, to be good, or to gain approval; exhibits excessive shyness and reserve; gives up easily; wants to be dressed, washed, fed, or fondled; seldom exhibits anger; and, is a follower. The third category of "internalizing behavior" included such behavior as restlessness in sleep, enuresis, embarrassment about toilet habits, nightmares, fearfulness, oversensitivity, perfectionistic tendencies, overconcern with cleanliness, nervousness, worries, and insufficient appetite. The fourth category of "externalizing behavior" consisted of destructiveness, displays of temper, defiance, cruelty, extreme domineering tendencies, and overactivity.

The critical ratio technique was used throughout to determine the significance of mean differences between the ratings for the control and experimental subjects on (a) specific behavior scores, (b) over-all adjustment scores, (c) relation between the severity of speech defect and the severity

of problem behavior, and (d) characteristic behavior patterns that might differentiate the two groups. In those cases where correlations were computed, the Pearson product-moment method was used. The correlations of the ratings of the two judges ranged from .70 to .89 with a median correlation of .78 in the nine behavior areas and .67 for the over-all adjustment rating. Neither judge had any consistent tendency to rate higher or lower than the other.

Results

Of the nine selected behavior areas studied, four areas, in addition to the over-all adjustment rating, were found to show differences between the experimental and control groups which were significant at the .01 level of confidence or better. From mothers' reports, the behavior areas in which the experimental group showed significant differences from their controls at the .01 level were sleeping, fears and anxieties, and tension; and, at the .001 level, in the area of peer relations and in over-all adjustment. No significant differences were found between the two groups of children in the areas of eating, toilet training, comfort patterns, or aggression.

In the area of sleeping behavior, the problems were centered around difficulties in falling asleep, restlessness, and other disturbances during sleep. In the area of fears and anxieties, the problems were centered around specific fears, areas of overconcern, unusual sensitivity, worries, withdrawal, and perfectionistic tendencies. In the area of tension, the children were considered by their mothers to be restless, sensitive, or nervous children who became upset and cried easily. They exhibited overt tensional outlets in such behavior as nailbiting, eyeblinking, pulling or twisting the hair, and in general overactivity. The experimental children were reported to have more disturbances in the area of peer relations. These problems were expressed in such behavior as a need to be first in games, a preference for children other than the child's own age or sex, an attempt to win other children's friendship by submissiveness or overgenerosity, a reluctance to stand up for one's self, and a preference for playing alone.

In addition to being rated as having poor "over-all adjustment" in comparison with the control children, certain personality traits were found to characterize the experimental children. The speech-defective group tended to be passive children who internalized their responses and were characterized by submissiveness, timidity, and a need for approval. Although there was a tendency for the severity of problem behavior to be related to the severity of speech involvement, the relation did not yield a significant level of confidence.

It is of interest to note that in no instance was the control group of normal-speaking children reported to have more problem behavior than the experimental group. They were reported as having similar problems, but to a lesser degree than the experimental children. However, it should also

be noted that a sizeable number of children of both groups were found to have only slight problems, or none at all, in many of the areas under study.

Discussion

The findings of this study tend to support the contention that articulatory speech problems of a functional nature are part of a larger picture of emotional and social behavior in a child's development. Based on reports by the mothers, the speech defects tended not to occur in isolation but frequently with other disturbances touching several areas of their social and personal adjustment, although this was by no means typical of every child in the experimental group.

The sleeping disturbance, the tension-reducing outlets, the fears, the inadequacies in peer relationships, plus the submissiveness and need for approval, suggest the possibility that more experimental children were operating under a greater degree of emotional stress than was true of their controls. A possible explanation for the existence of this stress may be hypothesized as being a generalized anxiety underlying the diverse symptomatology presented by many of the experimental children. If speech represents a form of self-assertion, then to "speak up" in the adult manner may be a frightening and forbidding prospect to apprehensive or submissive children who may have a need to display more infantile and nonassertive behavior in their particular environments. The speech defect in such cases may serve as a means of reducing anxiety and stress and reflect, as well, a fear of taking forward steps toward maturity and increasing independence.

Bearing in mind that similar problems were found with some frequency in both the experimental and control groups, it can be assumed that the dividing line between the behavior patterns and personality of children with speech problems of this type and children with normal speech remains a tenuous one, although statistically verifiable. The similarities between many of the children from both groups may be due to the fact that the experimental group contained children with varying degrees of speech involvement, some of which would be barely perceptible to the untrained ear. Another consideration that might add to the similarity of the two groups may be the fact that no child in our culture escapes entirely from some adjustive difficulty in his development toward emotional and social maturity. These factors may have tended to minimize behavioral differences between the two groups. Nevertheless, this study does suggest that more serious problem behavior of a particular kind, as well as certain specific personality characteristics, will be more apt to occur among children with functional defects of articulation than among a comparable group of normal-speaking children.

Because of the personality and adjustment problems of the speech handicapped child, rapport with the child, the development of self-confidence, and the creation of an atmosphere that encourages the child toward self-

expression and self-assertion may be therapeutically more valuable than the specific speech correction measures employed. Prognosis in any given case will depend more explicitly upon the personality and the approach of the speech therapist and upon the quality of the relationship established with the child than has heretofore been generally recognized in practice.

Summary

The purpose of the study was to test for a possible relationship between functional articulatory speech defects in children and personality and behavior patterns. Using an open-ended questionnaire type interview with mothers, the personality and behavioral background of a speech-defective experimental group and a normal-speaking control group was obtained in nine areas significant of child development and personality. The protocols were rated by two judges to obtain quantitative and qualitative differences in behavior, patterns of personality traits, and the correlation of speech and behavior problem severity.

From mothers' reports, the experimental group showed significant differences from their matched controls in their over-all adjustment, peer relations, sleeping behavior, manifestation of fears and anxieties, and tensional outlets. In personality, the experimental subjects tended to be passive children who internalized their responses and who were further characterized by submissiveness, timidity, and a need for approval.

Functional speech problems of this type, apparently, are not isolated phenomena but part of a total adjustive pattern. It is suggested that an underlying stress may be common to the diverse symptomatology here presented and that the infantile and nonassertive behavior could very well serve as anxiety-reducing devices to meet environmental pressure.

Therapeutically, rapport with the child, the development of self-confidence, and the creation of an atmosphere and relationship that encourages the child toward self-expression and self-assertion may be more valuable than the specific speech correction measures employed. In turn, these findings suggest a broadening of the speech therapist's training to include a firmer background in child psychology and development and in counseling and guidance techniques for both children and adults.

References

1. Allen, I. M. Defect of the speech function in childhood. *New Zealand med. J.*, 1947, 46, 297-307.
2. Beasely, J. Relationship of parental attitudes to development of speech problems. *J. Speech Hearing Disorders*, 1956, 21, 316-321.
3. Brown, F. W. Baby talkers. *Proc. Amer. Speech Corr. Ass.*, 1936, 7, 197-208.

4. Chapin, A. B., & Corcoran, M. A. A program for the speech-inhibited child. *J. Speech Hearing Disorders,* 1947, 12, 373-376.

5. FitzSimons, R. Developmental, psychosocial, and educational factors in children with nonorganic articulation problems. *Child Develpm.,* 1958, 29, 481-489.

6. Irwin, R. B. Speech comes to a five year old boy. *J. Speech Hearing Disorders,* 1946, 11, 197-203.

7. McCarthy, D. Language disorders and parent-child relationships. *J. Speech Hearing Disorders,* 1954, 19, 514-523.

8. Macfarlane, J. W., Allen, L., & Honzik, M. P. *A developmental study of the behavior problems of normal children between twenty-one months and fourteen years.* Univer. of California Press, 1954.

9. Mowrer, O. H. *Psychotherapy, theory and research.* Ronald Press, 1953.

10. Nelson, O. W. An investigation of certain factors relating to the nature of children with functional defects of articulation. *J. educ. Res.,* 1953, 47, 211-216.

11. Powers, M. H. Functional disorders of articulation—symptomatology and etiology. In L. E. Travis (Ed.), *Handbook of speech pathology.* Appleton-Century Crofts, 1957. Pp. 707-768.

12. Russell, C. M. Personality factors in a motor speech delay case. *Amer. J. ment. Def.,* 1944, 49, 171-176.

13. Spriestersbach, D. C. Research in articulation disorders and personality. *J. Speech Hearing Disorders,* 1956, 21, 329-335.

14. Van Alstyne, D., & Hattwick, L. A. Behavior patterns of children with nursery school experience. In R. G. Kuhlen & G. G. Thompson (Eds.), *Psychological studies of human development.* Appleton-Century-Crofts, 1952. Pp. 411-415.

15. Werner, L. S. Treatment of a child with delayed speech. *J. Speech Hearing Disorders,* 1945, 10, 329-334.

16. Wood, K. S. Parental maladjustment and functional articulatory defects in children. *J. Speech Hearing Disorders,* 1946, 11, 255-275.

17. American Speech & Hearing Association Committee on Midcentury White House Conference. Speech disorders and speech correction. *J. Speech Hearing Disorders,* 1952, 17, 129-137.

EFFECTS OF PRETRAINING ON THE REINFORCING VALUE OF VISUAL STIMULI[1]

HAROLD W. STEVENSON *and* RICHARD D. ODOM*
University of Minnesota

Although there have been numerous studies indicating that visual stimuli may operate as effective reinforcers for the behavior of subhuman species and human adults, there is little information concerning this problem with children. The present study seeks to investigate the effect of visual reinforcement on children's performance in a simple operant task. In addition, the question is raised whether visual reinforcers are subject to satiation and deprivation effects.

The effects of two pretraining conditions on performance are investigated. *Ss* in one pretraining condition were left alone for 15 minutes in a room equipped with a large variety of novel and interesting mechanical toys. In a second pretraining condition, *Ss* spent the 15 minutes alone in the room with no toys. It is assumed that the effectiveness of visual reinforcers would be reduced following experience with the toys because of a partial satiation of a need for visual stimulation. It is also assumed that isolation with no toys would increase such needs and would thereby increase the effectiveness of visual reinforcement.

METHOD

Subjects

Ss were 30 boys and 30 girls attending kindergarten, first and second grades of the Laboratory School of the University of Minnesota.[2] All children in attendance on the days of the study were used as *Ss*. *Ss* were above average in intelligence and socioeconomic status. Six *Ss* were discarded for reasons described below and were replaced by other *Ss*.

* Institute of Child Development, University of Minnesota, Minneapolis 14.

[1] This study was supported by a grant from the National Science Foundation (Grant G 9687) to the first author.

[2] The writers wish to express their thanks to the staff of the University Laboratory School for their cooperation in this study.

Apparatus

Experimental task. The apparatus consisted of a wooden frame 20 in. long and 14 in. deep upon which two plywood panels were mounted. Centered on the front of the vertical panel was a rectangular aperture into which a 2 in. high and 3 in. wide piece of milk glass was inserted. The second panel sloped from a front elevation of 3½ in. to a rear elevation of 6½ in. Centered on the sloping panel was a vertical wooden lever, 2½ in. high and ¾ in wide. The apparatus was painted a flat gray with the exception of the horizontal panel and lever which were painted flat black.

Behind the apparatus was a 16 mm. Dunning Animatic filmstrip projector which projected a 1¾ in. by 2 in. picture on the milk glass screen. The apparatus and the projector were placed on a small table in a 5½ ft. square room. A dark green screen 6 ft. high and 2 ft. deep bordered the rear of the apparatus, preventing *S* from seeing the projector and exploring the rear of the apparatus.

A mirror was mounted above and behind *S*'s chair so that *S*'s performance could be viewed from the adjacent recording room. A child's chair completed the furnishings of the experimental room.

A film strip containing colored pictures of common animals was used to provide visual reinforcement. The film strip contained a total of 35 different animals arranged in two random sequences. Separating each picture was a blank frame. Each picture was illuminated for 1 sec., and between pictures the screen was a uniform gray. The pictures were presented according to a variable ratio schedule with a mean of one picture for every 20 responses.

The presentation of stimuli and recording of responses were controlled by electronic programming and recording equipment located in a room to the rear of the experimental room.

Pretraining task. The pretraining room was 13½ ft. long and 6½ ft. wide and was located next to the experimental room. It contained two tables and four chairs. A one-way vision mirror was located in a door at the rear of the room. On the left wall was a 4 sq. in. speaker and a window covered by a curtain.

For *Ss* in the Toy condition the tables and chairs were covered with toys. The toys included a drumming clown, a drinking rabbit, a rocking-telephoning bear, a balloon-blowing bear, a typewriting doll, a bouncing bird, a butterfly-chasing duck, a spiral-shooting gun, a large stuffed bear, a stuffed rabbit, a beach ball, and a xylophone. All except the last seven toys were powered by batteries and were equipped with 2 ft. extension cords equipped with switches which could be easily operated by the children.

Procedure

Ss were randomly divided into three groups with an equal number of boys and girls in each group. For the two pretraining groups a 15-minute

period in the pretraining room preceded the experimental task. In the Toy group the pretraining room contained the toys described above. In the Isolation group no toys were present. The preliminary 15-minute period was omitted for the control group.

Ss were obtained individually by *E* from their classroom. While *E* escorted *S* to the laboratory room, *E* attempted to gain rapport with *S* by talking about school and related matters.

Pretraining task. For *Ss* in the Toy and Isolation groups *E* approached the experimental room, stopped, and told *S*, "Now the game is in that room right there, but I have to get it ready. You wait in here and I will come and get you when the game is ready." *E* walked with *S* into the pretraining room.

In the Isolation condition, *E* told *S*, "As soon as I get the game ready I will come back and get you. You wait in here until I get the game ready." *E* went out of the room leaving the door partly open.

In the Toy condition *E* told *S*, "I have a lot of toys in here and you can play with them while you are waiting. Play with any of the toys you want to and have fun." Again, *E* left the door partly open.

In both the Toy and Isolation conditions instrumental music was played softly through the speaker. This was introduced in an effort to get *S* to relax in the pretraining room and to reduce possible concern *S* might have in playing with the mechanical toys.

As *E* left the pretraining room, he started a stopwatch and went to a hallway from which he could observe *S*'s behavior through the one-way mirror of the pretraining room. *E* was thus able to intercede if S appeared to be disturbed or tried to leave the room during the 15-minute period. Three *Ss*, all in the Toy group, were discarded because their pretraining was discontinued when they showed evidence of having to go to the toilet. Three other *Ss* had to be reassured once by *E* but then continued pretraining with no further difficulties and were retained as *Ss*.

At the conclusion of the 15-minute period *E* returned to the pretraining room and told *S*, "The game is ready. You can come with me now."

Experimental task. *E* took *S* to the experimental room, seated *S* before the apparatus, and said, "This is a game we call the Television Game. See this handle? You pull it down and pictures will come on up here. Now you pull the handle and you can see the pictures." *E* waited until one picture was projected. If *S* stopped pulling the lever before this occurred, *E* urged him to continue pulling. After *S* had seen the first picture, *E* said, "O.K. I would like for you to keep pulling. I have to go to the other room and work the machine. I'll be back when the game is over." *E* then went into the adjacent recording room, where he monitored the electronic apparatus and observed *S*. The experimental task was concluded after seven minutes.

Ss in the Control group were taken directly to the experimental room, and the procedure for the experimental task was identical to that for the other two groups.

Three *Ss* were discarded during their performance on the experimental task; for one *S* the equipment failed, one *S* showed motor impairment, and one *S* verbalized that she thought that she was to remember the name of each picture and responded at a rate slower than any other *S* in the study.

Results and Discussion

The major results of the study are summarized in Figure 1, which presents the average number of responses made during each minute of the task in each group. The data are presented separately for boys and girls, since preliminary studies with visual reinforcement indicated the possibilities of a sex difference in performance. In both the Isolation and Control groups a rapid increase in the number of responses is evident in the first few minutes; by the seventh minute there is a general decrease in the number of responses. In both groups the number of responses made by boys after the first minute is consistently above that of the girls. In the Toy group performance begins at a high level, continues in a somewhat variable fashion, and begins to decline toward the end of the task. The clear sex differences found in the other two groups do not appear.

The significance of the results can be assessed more adequately by means of an analysis of variance incorporating an analysis of trend (1). This anal-

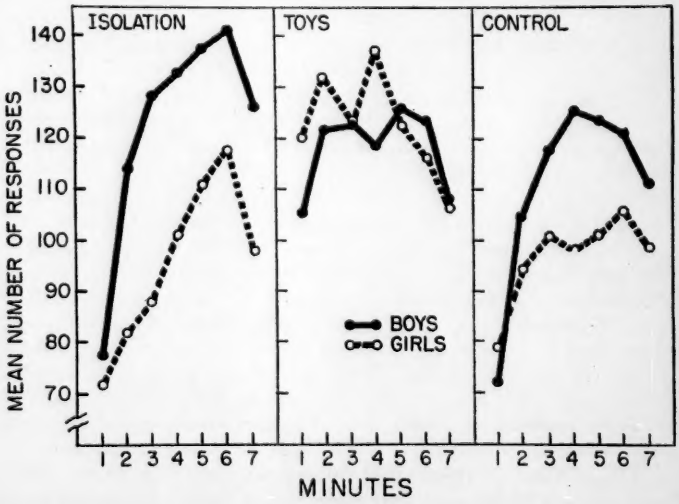

Figure 1—Average number of responses made by boys and girls in each group during the seven minutes of the experimental task.

TABLE 1

SUMMARY OF TREND ANALYSIS

Source of Variation	*df*	*MS*	*Error*	*F*
A. Over-all Trend	6	8648.8	E	20.9**
1. Linear	1	21564.5	E.1	17.6**
2. Quadratic	1	27669.1	E.2	48.0**
3. Residual	4	664.7	E.3	3.1*
B. Between Group Means	5	10890.1	D	2.5*
1. Sex (S)	1	13760.0	D	3.2
2. Condition (C)	2	10740.6	D	2.5
3. S × C	2	9604.6	D	2.2
C. Between Group Trends	30	913.3	E	2.2**
1. Linear	5	3652.4	E.1	3.0*
a. Sex	1	1634.3	E.1	1.3
b. Condition	2	8254.7	E.1	6.7**
c. S × C	2	59.3	E.1	. .
2. Quadratic	5	922.4	E.2	1.6
a. Sex	1	2564.3	E.2	4.4*
b. Condition	2	385.0	E.2	. .
c. S × C	2	639.0	E.2	1.1
3. Residual	20	226.2	E.3	1.3
D. Between Individual Means	54	4317.1	E	10.4**
E. Between Individual Trends	324	413.9		
1. Linear	54	1225.0		
2. Quadratic	54	576.7		
3. Residual	216	170.4		

* $p < .05$.
** $p < .01$.

ysis is summarized in Table 1. The linear and quadratic components of trend shown in Figure 1 are highly significant, indicating a general increase in response rate with a curvilinear trend being produced by the decline in response rate toward the end of the seven-minute period.

The *F* related to between-group means is significant, but the over-all performance of boys and girls and the over-all performance in the three conditions do not differ significantly. The effect of each variable on performance is evident, however, when the between-group trends are analyzed. The difference in quadratic trend is significant for the two sexes, and the difference in linear trend is significant for the three conditions. There are significant individual differences in average performance.

The data offer some support for the prediction that a 15-minute period of visual stimulation would reduce the reinforcing effects of visual stimuli. The rapid increase in response rates found in the Isolation and Control groups did not occur in the Toy group. *Ss* in the Toy group began response at a high level and maintained this level throughout most of the seven-minute period. There was no indication that visual reinforcement increased the rate of response for these *Ss*. The question posed by the data, however, is why the initial rate of response was so high. Two interpretations may

be offered: First, the pretraining experience may have increased the strength of the general class of manipulatory responses so that these *Ss* entered the experimental task with a strong tendency to respond. A second interpretation is that the pretraining experience had an activating effect which carried over into the experimental task to produce a high level of responding.

Isolation for a 15-minute period did not have a strong effect on performance. The rate of response of boys was consistently higher in the Isolation group than in the Control group, but the rate of response did not differ consistently between the two conditions for girls. One of the problems encountered in this type of study is to provide a meaningful degree of visual deprivation. It was hoped that the lack of novel and interesting objects would result in a higher degree of visual deprivation in the Isolation group than in either of the other groups; however, *Ss* had an ample opportunity to explore the room and gain visual experience during the 15 minutes of isolation.

The tendency of boys in the Control and Isolation groups to show a greater increase in response rates than the girls merits further investigation. It is doubtful that the difference is attributable to differences in motor ability, for the initial rates of response were very similar in all cases for boys and girls, and the general level of response of girls in the Toy group was approximately equal to that of the boys. There is the possibility, of course, that these particular pictures were more appealing to boys. Again, this does not seem to be a likely interpretation.

Summary

Three groups of kindergarten and elementary school children were presented with a simple operant task in which lever pressing was reinforced with visual stimuli. A variable-ratio schedule of reinforcement provided a colored picture of a common animal after an average of every 20 responses. A control group and two pretraining groups were included. The first pretraining group played alone for 15 minutes in a room containing a wide variety of interesting toys. The second pretraining group was left alone for 15 minutes with no toys. An analysis of trend of performance over the seven-minute experimental period indicated a significant difference in linear trend associated with condition. The Toy group began response at a higher level than the other two groups, but failed to show the rapid increase in response rate found in the other groups. The difference in the quadratic component of between-group trends associated with sex was significant.

References

1. Grant, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, 53, 141-154.

EFFECTS OF PRETRAINING WITH VARIED STIMULI ON CHILDREN'S DISCRIMINATION LEARNING[1]

SHELDON H. WHITE*
University of Chicago

It has been found that a response generalizes more when it has been learned to several stimuli than when trained to a single stimulus (4, 7). If varied training somehow induces *S* to minimize or overlook subsequent stimulus differences in the experimental setting, it should lead to poorer subsequent discrimination learning than equivalent pretraining to respond to a constant stimulus.

The first study to be reported sought to test this hypothesis by pretraining *S*s to respond to stimuli of constant or varied color prior to the learning of a two-choice color discrimination task. One group (C-8) was given eight trials of pretraining to choose the correct stimulus of a subsequent discrimination. A second group (V-24) was given the same training plus equal training to two irrelevant stimuli. A third group (C-24) was given 24 trials of pretraining to choose the correct stimulus.

A second study then sought to determine directly how varied pretraining transfers to discrimination learning, using two control groups. In this study, a group (V-24) given varied pretraining was compared with a group (N-24) given equivalent pretraining to a neutral stimulus and another group (O) which learned the discrimination problem without pretraining.

EXPERIMENT I

Method

Subjects. The *S*s were 60 3½- to 5-year-old preschool children, enrolled in two private Chicago nursery schools.[2] Twenty, 14 from school A and six from school B, were assigned by lot into each of the three experimental

* Department of Psychology, University of Chicago, Chicago 37, Illinois.

[1] Experiment 1 was performed with the aid of a Ford Foundation in Behavioral Sciences research grant, experiment 2 with the aid of a research grant (M-3639) from the National Institute of Mental Health, United States Public Health Service.

[2] The author wishes to express his gratitude to Mrs. Marguerite Kelly of the Bo-Peep Nursery School, Mrs. Mannie Tuteur of the Garden Valley Nursery School, Mrs. Dorothy Anker of the Home Club Nursery School, and Miss Jennie Lee Golding and Mrs. Esther Hecht of the South Shore Nursery School, together with their staffs, for their helpful and pleasant cooperation with this study.

groups. The children of school A were experimentally naive; about half of the school B children had participated in a previous study which involved discrimination learning.

Apparatus. The stimuli used in pretraining and discrimination were four 3½ in. squares, cut from ⅝ in. plywood and painted white, yellow, red-orange, and red—in Ostwald color notation: a, 3 pa, 6½ pc, and 7½ pc (1). These blocks were always shown on a 9 by 12 in. gray plywood base (Ostwald e). Holes in the base and underside of the blocks enabled a marble to be concealed under a block whether it was placed on the right or left side of the base. A flat-black wooden panel was placed between the child and the base to screen the *E*'s preparation for each trial.

Procedure. At the beginning of his session, the child chose one of a display of small toys to be his prize for playing the game. *E* took the toy and instructed the child that he was to fill the indentations on a marble board in order to win it as a prize.

During pretraining trials, *S* was faced with one block at a time and simply lifted it to find a marble underneath, which he saved towards his toy. *Ss* in group C-8 were given eight pretraining trials to the red-orange block. Group V-24, during 24 trials, had eight trials each to the white, yellow, and red-orange stimuli, with irregular alternation of the stimuli. Group C-24 had 24 trials to the red-orange block. For all groups, successive pretraining stimuli were alternated between the right and left sides in a pattern designed to avoid positional or single-alternation sequences of response.

The following instructions were given between pretraining and discrimination: "Now the game is going to get a little harder. I'm going to show you two blocks at a time, but only one of them is going to have a marble under it. Each time I show you the blocks, you have just one guess to try to find the marble—you can just pick up one block at a time. Try to get a marble each time." Following these instructions, the child was given 40 trials of a two-choice discrimination task, with the red-orange block correct and the red block incorrect. Counterbalancing for position was continued as before, with the correct stimulus appearing equally often on the left and right hand sides.

At the conclusion of the experiment, the child was always awarded his toy and encouraged to feel that he had done well.

Results

Despite the simplicity and repetitiveness of the pretraining procedure, all children responded well during pretraining and, without any obvious flagging of interest, collected their marbles and saved them towards their toy.

Upon transition to the discrimination task, they did not automatically continue with that familiar stimulus which had been bringing them reinforcement. Of the 20 *Ss* in each group, only about half gave a correct response on the first discrimination trial—10 in group C-24, 11 in group

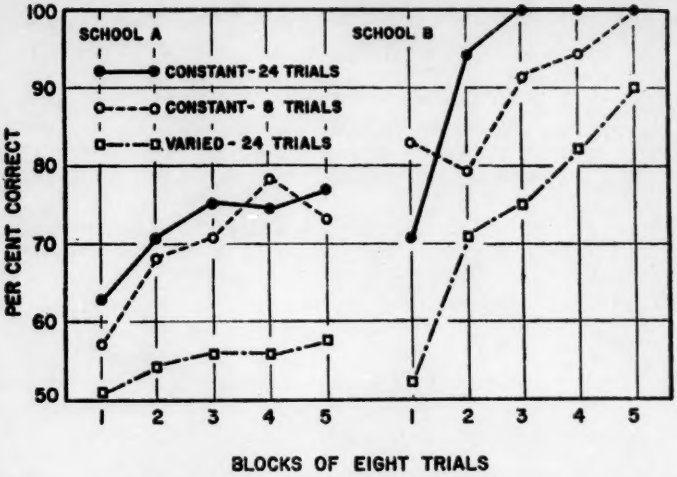

FIGURE 1—Discrimination learning after two degrees of pretraining to choose the correct stimulus (C-24 and C-8) and varied pretraining including the correct and two irrelevant stimuli (V-24). There were 14 children in each of school A groups, and six in each of school B groups (experiment 1).

C-8, and nine in group V-24. In this connection, it is possible that the transitional instructions ("Now the game is going to get a little harder . . .") caused them to take a fresh approach towards the discrimination task.

Complete performances of the three groups on the discrimination task are shown in Figure 1. Despite the fact that the *Ss* of school B performed at an appreciably higher level than those of school A, the information conveyed by either set of subgroups appears about the same. All groups showed some tendency to improve their performance over trials. C-24 *Ss* performed best, with C-8 *Ss* only slightly poorer; the performance of both groups was markedly superior to that of group V-24.

An analysis of variance (2) compared children of different groups and schools for number of correct responses in successive eight-trial blocks of the discrimination learning task (Table 1). The significant Groups effect reflects the superiority of groups C-24 and C-8 over group V-24. These three groups averaged, respectively, 31.4, 30.2, and 24.3 correct responses in 40 trials. Comparisons by t tests showed group V-24's performance to be significantly poorer than those of groups C-24 and C-8 ($p < .01$ and $p < .05$); the C-24 and C-8 groups did not differ significantly from each other ($t < 1$).

The significant Schools effect shows that the superior performance of *Ss* in school B is statistically reliable. The significant Blocks effect reflects the increase over trials shown by all subgroups. The two interactions which

TABLE 1

SUMMARY OF EXPERIMENT 1—ANALYSIS OF VARIANCE COMPARING NUMBER OF CORRECT RESPONSES IN EIGHT-TRIAL BLOCKS OF CHILDREN OF DIFFERENT GROUPS AND SCHOOLS

Source	*df*	*MS*	*F*	*p*
Between Subjects	59			
Groups (G)	2	57.22	5.22	.01
Schools (S)	1	161.60	14.74	.001
G × S	2	0.09	< 1.00	
error (b)	54	10.96		
Within Subjects	240			
Blocks (B)	4	17.99	12.94	.001
G × B	8	0.44	< 1.00	
S × B	4	2.86	2.06	.10
G × S × B	8	1.76	1.27	.10
error (w)	216	1.39		
Total	299			

approach significance in the table probably derive from the following trends (Figure 1): *Ss* in school B showed a steeper increase over trials than those in school A, which contrast between schools is relatively more marked for group V-24 than for the other two groups.

EXPERIMENT 2[3]

The first experiment showed that varied pretraining leads to poorer discrimination learning than constant pretraining. Is the varied pretraining less facilitating than the direct pretraining, or does it actually interfere with a subsequent discrimination? The second experiment sought an answer to this question by a direct assessment of the effects of varied pretraining on discrimination.

Method

The *Ss* were 60 3½- to 5-year-old children, experimentally naive, drawn from two additional Chicago nursery schools. Twenty, 10 from school C and 10 from school D, were assigned by lot into each of the three experimental groups.

The varied pretraining of group V-24 has been described above. Group O had no pretraining and was intended as an absolute control, whose com-

[3] Gratitude is here expressed to Mr. Paul Grim, who collected the data and assisted in its analysis for this experiment.

parison with group V-24 would show whether varied pretraining has a net effect of facilitating or interfering with discrimination learning. A second control group was considered necessary because of the generally facilitative effect which *any* sort of pretraining might be expected to produce, by "warm-up." Thus, group N-24 received 24 trials of pretraining to a neutral, gray (Ostwald 15 ml) block before going on to the discrimination. It was intended as a relative control, whose comparison with group V-24 would show whether varied pretraining produces more or less facilitation than can be attributed to its warming-up effect.

The discrimination task and other aspects of the method and procedure were as described above for experiment 1.

Results

In this experiment, the first trial of the discrimination showed choices biased towards the red-orange, correct stimulus. Forty-seven of the 60 *Ss* chose it, a proportion far beyond chance (binomial $p < .0001$). Of the 47, 14 were in group V-24, 16 in group N-24, and 17 in group O. On the next three trials, the red-orange stimulus was chosen 26, 21, and 36 times out of a possible 60. The most plausible interpretation of this choice pattern was in terms of the standard procedure of stimulus placement, with a right-position preference strong on the first trial but fading over the next few trials.

Table 2 shows the mean number of correct responses over 40 trials for each of the three groups. Although there are apparent differences among

TABLE 2

DISCRIMINATION PERFORMANCE (MEAN NUMBER CORRECT IN 40 TRIALS) FOR THE THREE GROUPS OF EXPERIMENT 2

	Group N-24	*Group V-24*	*Group O*	*All Groups*
School C	27.7	24.8	22.9	25.1
School D	26.1	26.5	21.1	24.6
Both Schools	26.9	25.6	22.0	24.8

the groups—principally, a greater number of correct responses for groups N-24 and V-24 than for group O—these differences were not confirmed in over-all analysis of the experiment 2 data. Analysis of variance with Groups, Schools, and Blocks (of eight trials) as factors showed only one significant entry: the Blocks effect, reflecting a reliable general increase in performance over successive blocks of trials.

COMBINED ANALYSIS OF EXPERIMENTS 1 AND 2

In each of the two experiments, the replications in different schools (school A vs. school B in experiment 1; school C vs. school D in experiment 2) generally confirmed each other, but were by no means numerically

identical in result. As was indicated by the significant Schools effect in experiment 1, such differences between replications do not necessarily come from sampling error but may reflect genuine population differences. For this reason, some caution must be exercised in comparing experimental groups drawn from different sample sources.

Some justification for cross-comparisons between experiments 1 and 2 may be found in the fact that their identically-treated V-24 groups gave mean performances, respectively, of 24.3 and 25.6 correct responses in 40 trials—with the difference between them being statistically and practically not significant. For this reason, and because there is a degree of "face validity" about the results, the cross-experiment comparisons are to be described. Conclusions arising from these comparisons depend upon the comparability of the experimental samples, and this reservation must be emphasized.

The first column of Table 3 shows the means of the five different groups obtained in the two experiments (with both V-24 samples pooled). A simple

TABLE 3

PERFORMANCE OF THE TREATMENT GROUPS OF EXPERIMENTS 1 AND 2 AND *t* TEST COMPARISONS

Group	*Mean*	*SD*	*N*	*t* TEST COMPARISONS			
				C-24	*C-8*	*N-24*	*V-24*
C-24 ...	31.4	8.57	20				
C-8	30.2	9.19	20	0.41			
N-24 ...	26.9	8.93	20	1.61	1.15		
V-24	24.9	7.41	40	2.87**	2.24*	0.86	
O	22.0	3.96	20	4.43**	3.66**	2.24*	1.97

* Significant at .05 level.
** Significant at .01 level.

analysis of variance, comparing the five groups, yielded an F of 5.12 (4,115 *df*), significant at the .001 level of confidence.

The right half of Table 3 shows the results of t test comparisons among the groups. Because the variance of the group O scores was markedly less than the variance of the other groups, all significance levels are for degrees of freedom determined by a formula which takes into account inequality of variances (6).

As one might expect, direct training to the correct stimulus is followed by the best discrimination performance, although it is perhaps somewhat surprising that 24 trials of such pretraining has little more effect than does eight trials. There is a noticeable superiority of this sort of training over a simple "warming-up" sort of pretraining (group N-24); the statistics do not show this superiority to be reliable, but, nevertheless, the hypothesis

probably might still be entertained that the facilitating effects of C-24 and C-8 pretraining are due to something more than simple warm-up.

That warm-up itself facilitates discrimination learning is shown by the significant superiority of group N-24 to group O. Group V-24 lies between groups N-24 and O, superior to O by a t significant only at the .10 level of confidence. Consequently, the interpretation of this kind of pretraining must be reserved; it probably is facilitative, although perhaps less than one might expect from its warm-up effect alone. Varied pretraining is certainly less facilitating than direct pretraining, even to that amount of direct pretraining (eight trials) which it itself includes.

Discussion

In experiment 1, the rather surprising finding that group C-24 was not markedly superior to group C-8 was noted, implying that 24 trials of direct pretraining to the correct stimulus is hardly more helpful to a subsequent discrimination than only eight such trials. Since the performance of both groups was at a rather high level, there is the possibility that group C-24 was hindered by a "ceiling effect" from expressing a potentially stronger degreee of superiority. This possibility cannot be fully discounted; however, it will be noted in Figure 1 that the C-24 – C-8 difference was about the same for schools A and B (quantitatively, about 4 per cent in each school) despite the fact that school B *Ss* were visibly operating much closer to ceiling than school A *Ss*. There may, really, be little difference in the effectiveness of these two degrees of direct pretraining.

The differences between the schools of experiment 1 may have come from several causes. Some of school B *S*'s were participants in an earlier discrimination study, and Shepard (5) has demonstrated a learning-to-learn effect for children's discriminations which is particularly potent from the first to the second problem. School B staff participated more, and was more supportive, in the recruitment of children for the experiment. Finally, there is the possibility of a relatively subtle socioeconomic difference: both schools enroll their children from substantially the same neighborhood, but school B is reputed to attract relatively more professional and educated people.

It was the author's impression that the two nursery schools of the second experiment differed more in character of administration and socioeconomic sample than did those of experiment 1, yet no strong difference between schools C and D was found.

Group N-24 of the second experiment requires some interpretation. As noted above, it was included as a measure of the general warming-up effects of single-stimulus pretraining. Facilitating transfer from warm-up is a well-substantiated effect in adult human learning (3) and is probably equally effective in children's learning, although there does not appear to be a direct experimental demonstration to this effect. Its effects upon adult learning have been attributed to set factors, particularly "attentional" and postural

adjustments to the experimental task. In the case of children, particularly where they are experimentally naive, it seems probable that there is an additional, emotional component of the child's warming-up to the situation —the overcoming of shyness before a relatively strange adult and his relatively strange new task. It might be an interesting question for further study to examine whether sophisticated preschool children would show quite as much benefit from pretraining as do naive children.

It appears that four of the pretraining groups—C-24, C-8, N-24, and O—arrange themselves in an order which is quite consistent with expectations from established mechanisms. Presumably the two best-performing groups would be superior through a combination of associative transfer of the habit of responding to the red-orange stimulus, plus nonspecific facilitating transfer from warmup. The lesser facilitation produced by N-24 pretraining would be based on warmup alone.

Group V-24, like group C-8, had eight trials of training to respond to the subsequently correct stimulus, but its additional training to irrelevant stimuli was apparently deleterious. Apparently the direct training is "lost" amongst the irrelevant. The performance of this group was below what one might have expected by warm-up alone, though not reliably so.

That varied pretraining has produced more stimulus generalization than constant pretraining is explainable either by the greater resistance to extinction which some studies have found after varied pretraining, or else by a "generalization set" (7). Since the influence of resistance to extinction is not so easily distinguishable in the present discrimination paradigm as it is in most generalization procedures, it is tempting to explain these results by the generalization set mechanism—with the varied pretraining serving as an "implicit instruction" that the response is to be widely attached or that stimulus differences may safely be overlooked.

Summary

In experiment 1, three groups of 20 preschool children each, drawn from two nursery schools, were given a pretraining procedure in which they saved marbles found under a single colored block. Group C-24 was presented a block of red-orange color 24 times, group C-8 the red-orange block eight times. Group V-24 was presented blocks of varied color—the red-orange, a yellow, and a white block—eight times each. Following their pretraining, all *Ss* were given 40 trials of a simultaneous color discrimination with the red-orange block correct and a red block incorrect. Groups C-24 and C-8 did not differ significantly from each other in discrimination performance, but were significantly superior to group V-24. Although the level of discrimination performance differed significantly between nursery schools, the essential experimental comparisons between groups were the same for both schools.

In experiment 2, there were also three pretraining groups of 20 preschool children each, who subsequently learned the same discrimination. In addition to group V-24, described above, group N-24 had 24 trials of pretraining to a neutral gray stimulus, and group O had no pretraining. Group N-24 was significantly superior to group O in discrimination learning, with group V-24 in between, not significantly different from either. Although the children were drawn from two nurseries, there were no schools differences.

Cross-comparisons between the two experiments ranked the five treatments in the following order in promoting discrimination: C-24, C-8, N-24, V-24, and O.

References

1. Jacobson, E., Granville, W. C., & Foss, C. E. *Color harmony manual.* (3d Ed.) Container Corp. of America, 1948.
2. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Houghton Mifflin, 1953.
3. McGeoch, J. A., & Irion, A. L. *The psychology of human learning.* Longmans, Green, 1952.
4. Riopelle, A. J., & Rogers, J. P. Transfer of training to orthogonal dimensions. *J. exp. Psychol.,* 1956, 52, 367-370.
5. Shepard, W. O. Learning set in preschool children. *J. comp. physiol. Psychol.,* 1957, 50, 15-17.
6. Walker, H. M., & Lev, J. *Statistical inference.* Holt, 1953.
7. White, S. H., & Spiker, C. C. The effects of a variable conditioned stimulus upon the generalization of an instrumental response. *Child Develpm.,* 1960, 31, 313-319.

SECONDARY REINFORCEMENT WITH PRESCHOOL CHILDREN[1,2]

JANE G. FORT*

University of Massachusetts

Although there is abundant literature in the area of secondary reinforcement, little, if any, of it can be considered conclusive evidence that stimuli can become and remain strong secondary reinforcers. In a recent review of the literature on secondary reinforcement (13), contradictory results obtained by Melching (12) and by N. A. Myers (14) were pointed out. The resolution of this difference in results could be a significant step in clarifying findings and their interpretations.

Melching (12) attempted to find secondary reinforcement effects in bar-pressing behavior subsequent to 50 per cent neutral stimulus (buzz) presentation in training. One half of these *Ss* underwent 100 per cent buzz in the extinction period, while the other half was presented no stimulus. The results indicated no significant difference in number of extinction responses made by these groups. Melching explains his results through stimulus generalization. Since the percentage of responses on which the buzz was experienced in training was changed an equal amount for both groups, equivalent extinction responses could then be expected.

Myers (14) used a controlled operant situation to study partial and regular reinforcement. Children were trained to make a series of responses

* Department of Psychology, University of Massachusetts, Amherst.

1 This research was supported by funds provided by NIMH Grant M-2620.

2 This article is based on a thesis submitted in partial fulfillment of the requirements for the Master's degree, Department of Psychology, University of Massachusetts, 1960 (6). The author wishes to extend her appreciation to her committee members, Drs. R. E. Wyman, A. E. Goss, N. A. Myers, and especially J. L. Myers for their patience, cooperation, and helpful advice.

for a candy reward. Contrary to Melching's findings, Myers found that, for each of her groups trained identically, the one given the newly acquired secondary reinforcer (chip) in extinction consistently made more responses in extinction than the one given no chip.

The discrepancy between these results and those of Melching may be due to one or a combination of factors. Melching employed rats in a free operant situation, using, for the potential secondary reinforcer, a buzz presented by the *E* at the same time as the primary reward. Myers' children had to manipulate a chip in order to receive their primary reward. It is possible that this handling enabled the chip to acquire greater reward value than did the mere hearing of the buzz used in Melching's study. However, species differences could be great enough to cause the cited discrepancy. The present study, therefore, was primarily designed to help clarify the results of the Melching and Myers studies. In addition, data were provided on partial primary reinforcement effects with children.

Method

Subjects

The *S*s were 144 preschool children ranging in age from 3 years to 5 years, 11 months.[3] Within the restrictions of equating the groups for age and for sex, the *S*s were equally and randomly divided among the 12 groups.

Apparatus

The apparatus, designed to attract and maintain the interest of preschool children, was a clown face, having red jewel-light eyes, a push-button nose, and a slot-tray mouth. Candy[4] (primary reinforcer) was dispensed through a tube to the mouth of the clown, while a 1-second buzz (potential secondary reinforcer) was heard from the interior of the clown. *E* had access to and operated two silent knife switches which allowed administration of the predetermined reinforcement. The number of responses was recorded on an electric magnetic counter mounted on the back of the clown, out of *S*'s sight. *E* recorded the cumulative number of responses made during each successive 30 seconds of extinction.

Procedure

A room separate from general distractions was chosen for the experimental session. The clown was placed on a small table or on the floor so that *S* had easy access to the nose. Subsequent to miscellaneous conversation

[3] The author extends her appreciation to the many Amherst, Massachusetts nursery schools and kindergartens and in particular to William Barry, Superintendent of Schools, and Miss Esther Wien, Elementary School Supervisor, Northampton, Massachusetts, for their cooperation in providing subjects and facilities for this study.

[4] The candies are manufactured by the Kirkman Pharmacal Company of Seattle, Washington. They contain no medication or sugar, but are sweet and come in four assorted colors and flavors: lime, lemon, orange, and raspberry.

designed to reduce possible tension, *S* accompanied *E* into the room. Attention was called to the clown face and especially to the nose. *S* was told that "something funny happens when you press his nose. Let's see what happens." *E* pressed the clown's nose and thereby received a candy. *S* was encouraged to try it also and was given one rewarded preliminary trial. He was then told that he could stay and continue to play the game "as long as you want." *S* and *E* remained alone in the room throughout the experimental session.

Each *S* was allowed 20 responses during training. According to a random and predetermined schedule for each group, *S* could experience one of four treatments following each response: (a) a candy alone, (b) a buzz alone, (c) a candy and a buzz, or (d) neither a candy nor a buzz. For each group, candy was presented with or without a buzz on the first and the last responses in training. On the next response after the 20 training responses were completed, the extinction period was noiselessly begun without comment from *E*. Each *S* was run until he stopped and indicated a desire to leave the room. If *S* had not stopped by the end of 10 minutes in extinction, *E* terminated the session. Questions and comments were ignored as much as possible, but if necessary, *E* made a very noncommittal response. To reduce frustration which might arise from the absence of the primary reward in extinction, each *S* was given one last candy at the end of the extinction period.

Experimental Design

During training, two schedules of primary reinforcement (100 per cent and 50 per cent candy) were used along with three schedules of potential secondary reinforcement (100 per cent, 50 per cent, and 0 per cent buzz). During extinction, two schedules of buzz (100 per cent and 0 per cent) were employed. There were 12 *Ss* in each of the resulting 12 groups.

Results

To reduce heterogeneity of variance, square root transformations were applied to the data. Analysis of variance was then run on the transformed total number of responses in extinction. Table 1 shows a summary of the analysis. Only the partial primary reinforcement effect was significant. The *F* was 15.05, yielding significance beyond the .001 level. Those *Ss* who chose to stop before the end of 10 minutes were given values of 0 for the remaining periods of time. Figure 1 shows the subsequent mean number of responses in extinction over 10 one-minute periods of time when the two percentages of primary reinforcement (100 per cent and 50 per cent) are plotted. The curves in both cases show a steep initial drop within the first two minutes and a subsequent gradual decrease to the same level. This figure gives a strong indication of the partial primary reinforcement effect. The groups which received 50 per cent candy presentation in training made a mean of 78 responses in the first minutes of extinction, while those trained

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE

Source	*df*	*MS*	*F*
Total	143		
Primary Reinforcement (P)	1	236.44	15.05**
Secondary Reinforcement (S)	2	2.52	.16
Extinction Reinforcement (E)	1	1.56	.09
P × S	2	20.52	1.30
P × E	1	28.62	1.82
S × E	2	11.99	.76
P × S × E	2	3.44	.21
Ss/P × S × E	132	15.71*	

* Error term for all sources of variance.

** $p < .001$.

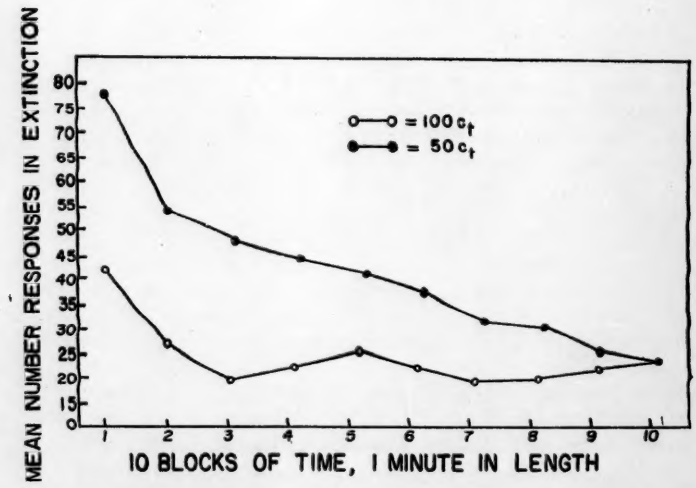

FIGURE 1—Mean number of responses in extinction in each of 10 one-minute periods of time for groups receiving 100 per cent and 50 per cent primary reinforcement in training.

TABLE 2

SUMMARY OF RESULTS FOR NUMBER OF RESPONSES IN EXTINCTION*

Group Number	Training Condition		Extinction Condition	Total Number Responses in Extinction
	Candy—%	Buzz—%	Buzz—%	
9	50	50	100	1174
11	50	0	100	1153
2	100	100	0	1015
10	50	50	0	843
8	50	100	0	833
7	50	100	100	657
12	50	0	0	589
6	100	0	0	522
1	100	100	100	480
5	100	0	100	440
4	100	50	0	331
3	100	50	100	307

* The groups may be referred to by the respective percentages of reinforcement administered in training and extinction; e.g., 100-50-0 denotes the group which received 100 per cent candy and 50 per cent buzz in training, and was presented 0 per cent buzz in extinction.

on 100 per cent primary reinforcement averaged 44 responses during this period.

Table 2 shows the total number of responses made by each group in extinction with the respective training and extinction conditions indicated. The groups are rank-ordered from the largest number of responses to the smallest. Here, it can be seen that with one exception (group 2: 100-100-0), the groups trained under the partial primary reinforcement schedule rank above those trained under the 100 per cent reinforcement schedule. The column which notes the percentage of responses on which the buzz was presented in training shows that the groups are not orderly with respect to this variable, but seem to be randomly distributed. Although the highest group in number of extinction responses was one which received the buzz on 50 per cent of the responses in training, two other groups which were also trained under this schedule (group 3: 100-50-100, and group 4: 100-50-0) were the lowest groups of all in number of responses in extinction. It may be noticed that the second highest group was one having never experienced the buzz in training and had it introduced in extinction (group 11: 50-0-100). The same random distribution noted in training is also noted for the percentage of responses on which the buzz was presented in extinction. Aside from the fact that the two highest groups were presented 100 per cent buzz in extinction, there is no regularity shown for this variable. Table 3

TABLE 3

SUMMARY OF RESULTS FOR MEANS AND VARIANCES

	CONDITIONS DURING EXTINCTION			
	100% BUZZ		0% BUZZ	
Conditions During Training	*Means*	*Variances*	*Means*	*Variances*
100% *Candy During Training*				
100% buzz	40.0	3,506.0	84.6	15,546.0
50% buzz	25.6	2,469.0	27.6	1,026.0
0% buzz	36.7	2,780.0	43.5	2,298.0
50% *Candy During Training*				
100% buzz	54.8	1,123.0	69.4	2,298.0
50% buzz	97.8	11,924.0	70.2	2,554.0
0% buzz	96.1	8,750.0	49.1	1,737.0

shows the means and variances for the 12 groups. Here, the heterogeneity of variance can easily be seen.

A Duncan range test (4) performed on the mean number of responses for each group showed that the two highest groups in number of extinction responses were significantly different from the two lowest groups. Group 9 (50-50-100) and group 11 made significantly more responses in extinction than did either group 3 or group 4.

DISCUSSION

In most studies of partial primary reinforcement, *Ss* have been rats or adults (8, 10); children have been used in only a few. Lewis (9) used 6- and 7-year-old boys in a gambling situation with toys as the primary reward. *Ss* were told that they could win more toys than they had been given, but that they could also lose all of those given them. Partial reinforcement percentages of 50 and 60 per cent produced greater resistance to extinction than 100 and 0 per cent reinforcement. Fattu, Mech, and Auble (5) trained 3- to 5-year-old children to pull a lever for candy. They found a group given 25 per cent reinforcement made more responses during 14 minutes of extinction and took longer to extinguish than did either 50 or 100 per cent reinforcement groups. Myers (14) also found a significantly greater number of responses in extinction following 50 per cent than following 100 per cent reinforcement. Thus, the results of the present study agree with those of other studies in which *Ss* were children.

In view of results found with partial primary reinforcement, one might anticipate that partial secondary reinforcement groups would exhibit a

similar effect. However, there was no indication of a secondary reinforcement effect in the present study. Number of responses made in extinction was not related to percentage of secondary reward in training. It was also noted that *Ss* presented the buzz on all responses in extinction did not make significantly more responses than those to whom the buzz was not presented.

The absence of a secondary reinforcement effect in the present study may have been due to several factors. First, there is the possibility of aversive properties of the buzz. However, if the buzz is a noxious stimulus, the groups which did not experience the buzz during training and then had it presented in extinction should have made fewer responses in extinction than those which did experience the buzz in training. This, however, is not the case, for the second highest group (group 11: 50-0-100) in terms of number of responses in extinction is a group which never experienced the buzz in training, but which had it suddenly introduced with the onset of extinction. Although the buzz may not be an aversive stimulus, it has been pointed out elsewhere (13) that stimuli are not really "neutral" and may not have equal initial reinforcing properties. The results of the present study would tend to support this idea. Some *Ss* in this experiment were trained under experimental conditions similar to those used by Myers (14). One difference is that the stimulus employed as the potential secondary reinforcer was a buzz in the present study, but a chip in the Myers study. It would seem from comparisons of groups which received similar treatments in these two studies that the chip has greater reinforcing potential than does the buzz.

Second, there is the possibility that the candy used in this study was not sufficiently reinforcing in and of itself to enhance the value of the buzz. Wyckoff (17) has suggested the need for very strong primary reinforcers in order to obtain an appreciable secondary reinforcement effect. However, there is no particular reason to doubt the rewarding properties of the candy; no *Ss* reported distaste for it, and it has been previously used successfully (2). There does remain a definite need for thorough research on various kinds of reinforcers and evaluation of their effectiveness. Valuable information may well be gained from actual scaling of stimuli proposed for use as reinforcers for children.

Third, the absence of a secondary reinforcement effect could be a result of too few pairings of the primary reinforcer and the potential secondary reinforcer. A few studies have investigated the effects of number of pairings of primary reinforcement and neutral stimulus on secondary reinforcement. Bersh (1) studied varying numbers of primary and neutral stimulus pairings and found that only his 20- and 120-pairing groups differed significantly from each other. This same general tendency was found by Hall (7), whose 75-reinforcement group was significantly different from his 25-reinforcement group. The Melching study found secondary reinforcement effects with 160 training trials. However, with children, Myers found that as few as 10 trials were sufficient to yield a significant secondary reinforcement effect.

Fourth, there is the possibility that the lack of a secondary reinforcement effect is the result of a great deal of *S* variability. It may be that much of the variance within groups is due to wide variation between *S*s in initial free operant levels. This, however, is doubtful, for the variation in the present study is not much different from that in a study in which strong secondary reinforcement effects was found (14). In addition, the ordering of the groups is so unsystematic as to suggest that the effect, if existing at all, is extremely slight and difficult to detect in situations employing children as *S*s and a buzz as the potential secondary reinforcer. However, the use of a preliminary measure by which pretraining operant level may be controlled for each treatment group should enable a clearer demonstration of the effects of the particular treatments employed.

None of the above factors, alone, then, seems sufficient to account for the lack of a secondary reinforcement effect in the present study. Very few studies have shown durable secondary reinforcement effects, and these studies have either employed special training schedules (e.g., 11, 15, 18) or a manipulable stimulus (e.g., 3, 14, 16). The previously-mentioned discrepancy, then, between the Melching and Myers studies would seem to be explained by differences in treatment of the potential secondary reinforcer. The children in the Myers study had to handle the chip and use it to receive the primary reinforcement, whereas Melching's rats had no physical contact with the buzz. It would seem that this physical contact with the chips led to much greater enhancement of their original value than was true of Melching's buzz or of the buzz in the present study. The results of the present study, then, would tend to indicate that mere pairing of a stimulus with a primary reinforcer is *not* a sufficient condition to establish noticeable and durable secondary reinforcement effects and that the mode of presentation and treatment of the potential reinforcer is of crucial importance.

Summary

The present study was concerned with reinforcement and extinction of a free operant response in preschool children. One hundred and forty-four *S*s were individually given an opportunity to "play a game" with a clown. When *S* pressed the nose, a candy fell into the clown's mouth. At the same time, a buzz was heard from his interior. Twelve experimental groups were used with both partial and regular schedules for the candy and the buzz. *S*s were given the predetermined reinforcement on the first 20 responses, which constituted the training period. Schedules of 100 per cent and 50 per cent candy and 100 per cent, 50 per cent, and 0 per cent buzz were used. In the extinction period, no candy was given and the buzz was presented on 100 per cent or 0 per cent of the responses until *S* indicated a desire to end the session. Otherwise, *E* terminated the session at the end of 10 minutes.

For each *S* the total number of responses in the extinction period was recorded. The results of this study indicated that:

1. Groups trained with 50 per cent candy reinforcement made more responses in the extinction period than did those trained with 100 per cent candy reinforcement.

2. There were no significant differences between groups as a result of the percentage of responses on which the buzz was presented in training or extinction. That is, there was no significant secondary reinforcement effect. Several possible explanations for this were discussed. It was concluded that treatment of the potential secondary reinforcer is important in the demonstration of noticeable and durable secondary reinforcement effects.

References

1. Bersh, P. J. The influence of two variables upon the establishment of a secondary reinforcer for operant responses. *J. exp. Psychol.*, 1951, 41, 62-73.
2. Bijou, S. W. A standard "candy" for research with children. *J. exp. Anal. Behav.*, 1958, 1, 314.
3. Cowles, J. T. Food-tokens as incentives for learning by chimpanzees. *Comp. Psychol. Monogr.*, 1937, 14, No. 5.
4. Duncan, D. B. A significance test for differences between ranked treatments in an analysis of variance. *Va. J. Sci.*, 1951, 2, 171-189.
5. Fattu, N. A., Mech, E. U., & Auble, D. Partial reinforcement related to "free" responding in extinction with pre-school children. *J. exp. Educ.*, 1955, 23, 365-368.
6. Fort, J. G. Extinction of a free-operant response in pre-school children following partial and regular schedules of primary and potential secondary reinforcement. Univer. of Mass. Library, 1960.
7. Hall, J. F. Studies in secondary reinforcement: I. Secondary reinforcement as a function of the frequency of primary reinforcement. *J. comp. physiol. Psychol.*, 1951, 44, 246-251.
8. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, 47, 193-234.
9. Lewis, D. J. Partial reinforcement in a gambling situation. *J. exp. Psychol.*, 1952, 43, 447-450.
10. Lewis, D. J. Partial reinforcement: a selective review of the literature since 1950. *Psychol. Bull.*, 1960, 57, 1-29.
11. McClelland, D. C., & McGown, D. R. The effect of variable food reinforcement on the strength of a secondary reward. *J. comp. physiol. Psychol.*, 1953, 46, 80-86.

12. MELCHING, W. H. The acquired reward value of an intermittently presented neutral stimulus. *J. comp. physiol. Psychol.*, 1954, 47, 370-374.

13. MYERS, J. L. Secondary reinforcement: a review of recent experimentation. *Psychol. Bull.*, 1958, 55, 284-301.

14. MYERS, N. A. Extinction following partial and continuous primary and secondary reinforcement. *J. exp. Psychol.*, 1960, 60, 172-179.

15. SALTZMAN, I. J. Maze learning in the absence of primary reinforcement: a study of secondary reinforcement. *J. comp. physiol. Psychol.*, 1949, 42, 161-173.

16. WOLFE, J. B. Effectiveness of token rewards for chimpanzees. *Comp. Psychol. Monogr.*, 1936, 12, No. 60.

17. WYCKOFF, L. B. Toward a quantitative theory of secondary reinforcement. *Psychol. Rev.*, 1959, 66, 68-78.

18. ZIMMERMAN, D. W. Durable secondary reinforcement: method and theory. *Psychol. Rev.*, 1957, 64, 373-383.

SECONDARY REINFORCEMENT AS A FUNCTION OF THE NUMBER OF REINFORCED TRIALS[1]

NANCY A. MYERS,* GRACE J. CRAIG,
and JEROME L. MYERS
University of Massachusetts

The purpose of this study was to observe the effects of varying habit strength sustained through secondary and primary reinforcers upon extinction behavior of children in an operant situation. A few studies (1, 3, 4, 6) have investigated the relation between number of training trials and extinction with secondary reinforcement of rats. Little research has been concerned with children's habit strength behavior. Lewis and Duncan (5) investigated the resistance to extinction of a lever-pulling response of college students after 3, 6, 12, or 21 acquisition plays with either 33 per cent, 67 per cent, or 100 per cent reward. Their results corroborated those of Capaldi with rats, indicating that the larger number of acquisition trials resulted in fewer plays to extinction. Siegel and Foshee (8) studied the law of primary reinforcement with preschool children, using 2, 4, 8, or 16 candy reinforcements for bar-pressing, and reported a positive relation between resistance to extinction and number of prior reinforcements. Spiker (9), however, found no significant increase in mean number of lever-pulling responses after 24 reinforcements than after 12 with preschool children, and Myers (7) found no differences in children's extinction behavior after 10 or 20 preceding candy reinforcements.

This study will evaluate the effects of number of training trials, with and without secondary reinforcement, on behavior of preschool children extin-

* Department of Psychology, University of Massachusetts, Amherst.

[1] This research was supported by funds from an NIMH Grant, M-2620.

Child Develpm., 1961, 32, 765-772.

guished with and without secondary reinforcement. Thus, it should help resolve the results of studies with rats and with humans, as well as extend the data on secondary reinforcement effects.

Method

Apparatus

The apparatus employed was a portable box designed to have eye-appeal and interest for children 3 to 6 years of age. It was placed on a low kindergarten table, and the child sat before it on a low chair. The front of the box was painted with the face of a clown. The back of the box was open for easy access to the candy and token reinforcement mechanisms, two counters, and three switches, by means of which *E* could turn *off* or *on* the lights, the token reinforcer, and the candy reinforcer. The clown's eyes were 6-watt bulbs which lit up at the start of each session and remained on throughout the session. The nose was a red cylinder protruding from the box, and on its end were two small push-buttons mounted horizontally parallel and an inch apart. Both the candy reinforcer and the token reinforcer delivered their contents to the mouth of the clown. A slot for inserting tokens was cut in one cheek, and a small box beneath this, on the inside of the apparatus, collected the tokens.

The token and candy dispensers were identical units differing only in the diameter of the loading tubes. They were designed to release an individual token or candy, respectively, to a slanted metal slide down which it traveled to the mouth of the clown. The loading tube for tokens permitted the stacking of 200 tokens; that for candy was tall enough to accommodate 30 candies. These tubes were refilled before each experimental session. The tokens used were small poker chips, about one half standard size and cut from four colors of plastic: red, white, blue, and yellow. These colors appeared randomly. The candy reinforcements were small chocolates with different colored shells, the popular M & M chocolates produced by Marlon & Marlon, New York.

Subjects

The *S*s were 48 boys and 48 girls between the ages of 5 years, 1 month, and 6 years, 4 months, attending kindergarten classes in the Easthampton, Massachusetts, Schools.[2]

Instructions

No written instructions were read to *S*, but *E* standardized the verbal instructions as much as possible. Each *S* was asked in the classroom if he would like to play a clown game. If *S* said yes, the teacher's permission was

[2] The authors wish to thank S. J. Bernard, Superintendent of Schools in Easthampton, Massachusetts, and his principals and teaching staff for their cooperation in permitting us access to the students and the school facilities.

obtained to leave the room with *E*. Once in the testing room *S* was instructed to sit at the small chair in front of the small table holding the box, and *E* would show him how to play the game.

The *S* was instructed through the following sequence: press the left push-button on the nose, receive a token, deliver the token to the slot in the cheek, push the right push-button on the nose, receive and eat the candy. This sequence was sufficient to establish the response in all *S*s. *E* then sat behind the box on the other side of the table, with easy access to the switches. As soon as the eyes were illuminated to signal the start of the game, *S* pressed the first button, and conditioning trials ensued. The eye-lights remained on continuously throughout conditioning and extinction. *S*s were told they could eat the candies, and in most cases consumed them at once. The only variations in the technique were for the four control groups (P-O groups) which never received tokens. They were instructed in the same manner, but were told that one button was not working, to push the second button only, and then receive a candy directly.

Design

All *S*s (except P-O groups *S*s) received one initial trial on the basic response sequence, in which they observed the lights, pressed the first nose-button, received a token, delivered it to the slot in the cheek, pushed the other nose-button, and received a candy. They were then assigned to the eight groups of eight *S*s each, four boys, four girls, for the remaining reinforcement and extinction procedures. The groups were equated for mean age. They differed in number of training trials and in extinction conditions. Two groups received two training trials, two groups received four training trials, two groups received eight trials, and two groups received 16 trials before extinction. One of each of the groups at each habit-strength value was then extinguished with *tokens* delivered when the token button was pressed (S^r-S^r groups); the other received no tokens for token-button pressing (S^r-O groups). No groups received candy for pressing the candy button in extinction. Two counters reported the cumulative number of responses made on each button, and *E* recorded the number of responses made on each button in each minute of the 5-minute extinction period.

P-O groups for which the token dispenser was not used during training received 2, 4, 8, or 16 trials with candy reinforcement and no candy for button-pressing in extinction.

Results

An analysis of variance was performed on the square-root transforms of the data of the eight experimental groups with respect to total number of extinction responses on both buttons. A summary of this analysis may be seen in Table 1. There were no differences in total number of extinction responses as a function of number of training trials. There were also no

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE PERFORMED ON TOTAL RESPONSES ON BOTH BUTTONS FOR EIGHT GROUPS (S^r-S^r AND S^r-O)

Source	*df*	*MS*	*F*
Number of Trials (Tr)	3	.98	.27
S^r	1	9.03	2.49
Tr × S^r	3	4.26	1.17
Ss/Tr × S^r	88	3.62	
Between *Ss*	95		
Button (B)	1	1.26	1.26
Time (T)	4	16.91	52.84**
B × T	4	.10	....
B × Tr	3	.78	.78
B × S^r	1	23.84	23.84**
T × Tr	12	.26	....
T × S^r	4	.19	....
B × S^r × Tr	3	.69	.69
T × S^r × Tr	12	.24	....
B × T × Tr	12	.26	1.52
B × T × S^r	4	4.83	28.41**
B × T × S^r × Tr	12	.15	....
Ss × B/Tr × S^r	88	1.00	
Ss × T/Tr × S^r	352	.32	
Ss × B × T/Tr × S^r	352	.17	
Within *Ss*	864		
Total	959		

** $p \leq .01$.

significant differences in total number of extinction responses for those groups given tokens in extinction and those who received nothing. All groups extinguished, i.e., there was a significant trend over time to fewer and fewer responses ($p < .01$). In addition, the Button × Secondary Reinforcement and the Button × Secondary Reinforcement × Time interactions were significant ($p < .01$), as may be seen in Figure 1. This figure shows the mean number of responses made on the token button and on the candy button during each minute of extinction for the four groups which received S^r in training and extinction (S^r-S^r) and for the four groups which received S^r during training trials, but none in extinction (S^r-O). The solid lines connect the mean number of responses to the *candy* button for the S^r-S^r (open circles) and S^r-O (closed circles) groups separately. The S^r-S^r groups made almost four times as many candy-button presses in the first minute of extinction as did the S^r-O groups, but made about the same number of

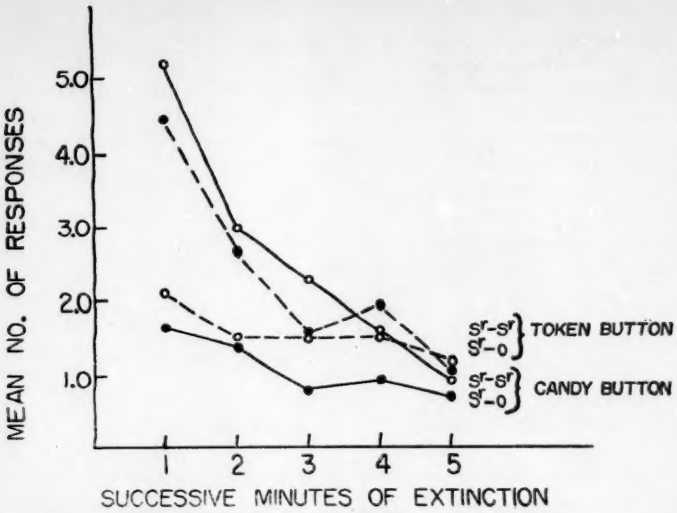

FIGURE 1—Mean number of responses on each button for each minute of extinction for the four E-groups receiving tokens in extinction (S^r–S^r) and for the four E-groups receiving nothing in extinction (S^r–0).

TABLE 2

NUMBER OF RESPONSES ON EACH BUTTON FOR EACH MINUTE OF EXTINCTION FOR S^r-S^r, S^r-O, AND P-O GROUPS

	MINUTES OF EXTINCTION					
	1	*2*	*3*	*4*	*5*	*Total*
S^r-S^r *Groups*						
On Token Button	101	72	73	74	58	378
On Candy Button	251	145	111	77	45	629
Total	352	217	184	151	103	1007
S^r-O *Groups*						
On Token Button	215	129	75	96	53	568
On Candy Button	82	65	38	45	31	261
Total	297	194	113	141	84	829
P-O *Groups*						
On Token Button	9	10	22	18	3	62
On Candy Button	293	208	209	174	161	1045
Total	302	218	231	192	164	1107

candy-button presses as the S^r-O groups in the last minute of extinction. The dotted lines in Figure 1 connect the mean number of responses made to the *token* button for the S^r-S^r (open circles) and S^r-O (closed circles) groups. It may be seen that the S^r-O groups made approximately twice as many token-button presses as the S^r-S^r groups in the first minute of extinction. The S^r-S^r groups made more candy-button responses and fewer token-button responses than did the S^r-O groups.

Table 2 presents a summary of the extinction data for the S^r-S^r, S^r-O, and P-O groups. An analysis comparing the total number of button presses of the experimental groups with the four groups (P-O) who received no S^r even during training revealed no differences in performance.

Discussion

The variation in number of trials before extinction did not produce a significant effect in the present experiment. A prediction of more extinction responses with increased numbers of training trials might ordinarily be made for a learning situation. The response in this study, however, was one well established in the repertoire of preschoolers, and the situation might be more accurately termed a performance one. Lewis and Duncan (5) first commented on this distinction and predicted, for an overlearned response, that increased training trials would decrease the number of extinction responses, due to a facilitation of discrimination between training and extinction. The number of training trials in this instance varied between two and 16; it is possible that 16 regularly reinforced trials were not more effective than two in enabling the *Ss* to discriminate extinction from training. All groups showed a very low level of responses in extinction.

Although the groups given tokens in extinction (S^r-S^r) made *more* responses in toto than did the groups extinguished with nothing (S^r-O), the difference merely approached significance in this study. However, it must be remembered that token reinforcement when administered actually produced a delay in button-responding, as it necessitated the operations of retrieving and inserting the token before another button-press was called for. Despite this enforced delay, the token reinforced groups made more responses.

The use of two buttons in this situation demonstrated the "chaining" function of the S^r. The significant Button $\times$ S^r interaction revealed that the groups given S^r in extinction made more candy-button presses and fewer token-button presses than did the S^r-O groups. The S^r-O groups received nothing for responding on the token button. This operated to break the established chain of behavior: The *Ss* tended, continually and immediately, to repress the token button, which led to a high initial rate of responding on the token button and to more token-button presses for the S^r-O groups than for the S^r-S^r groups. The number of token-button presses per

unit time decreased rapidly for the S^r-O groups, however, as seen in Figure 1.

The S^r, when administered in extinction, seemed to release the additional behavior in the chain, i.e., insertion of the token and subsequent pressing of the candy button. Thus, the S^r-S^r groups made more candy-button presses than did the S^r-O groups. Zimmerman (10) has suggested that this release of new behavior patterns is a necessary ingredient for strong S^r effects. In this particular experimental situation, at least, it was not a sufficient condition. However, the S^r seems to have had some negligible effect upon token-button performance. The rather constant rate of S^r-S^r groups' token-button pressing throughout the five-minute extinction period should be noted in Figure 1 and contrasted with the decreasing rate of token-button pressing revealed by the S^r-O groups. If the extinction period had been extended, this suggested difference in trend might have shown more distinctly.

Summary

This experiment was concerned with the extinction behavior of children following varying numbers of reinforced trials. Children were presented individually with a "clown game." When a button on the nose of the clown was pressed, a token was delivered from the clown's mouth. The token was inserted in the slot provided on the clown's face, and a second button on the nose pressed. This time a candy fell from the mouth. There were 12 groups of eight *Ss* each, differing in reinforcement and extinction procedures and equated for mean age and sex. Each child received either 2, 4, 8, or 16 training trials with either token *and* candy reinforcement, or candy reinforcement alone, followed by five minutes of extinction. There were no differences in total group performance due to the number of training trials or secondary reinforcement. More candy-button responses were made by groups extinguished under S^r; more token-button responses were made by the groups receiving nothing in extinction.

References

1. Bersh, P. J. The influence of two variables upon the establishment of a secondary reinforcer for operant responses. *J. exp. Psychol.*, 1951, 41, 62-73.
2. Capaldi, E. J. Effect of different amounts of alternating partial reinforcement on resistance to extinction. *Amer. J. Psychol.*, 1957, 70, 451-452.
3. Capaldi, E. J. Effect of different amounts of training on resistance to extinction of different patterns of partially reinforced responses. *J. comp. physiol. Psychol.*, 1958, 51, 367-371.

4. HALL, J. F. Studies in secondary reinforcement: I. Secondary reinforcement as a function of the frequency of primary reinforcement. *J. comp. physiol. Psychol.*, 1951, 44, 246-251.

5. LEWIS, D. J., & DUNCAN, C. D. Expectation and resistance to extinction of a lever-pulling response as a function of percent reinforcement and number of acquisition trials. *J. exp. Psychol.*, 1958, 55, 121-128.

6. MILES, R. C. Relative effectiveness of secondary reinforcement throughout deprivation and habit strength parameters. *J. comp. physiol. Psychol.*, 1956, 49, 126-130.

7. MYERS, N. A. Extinction following partial and continuous primary and secondary reinforcement. *J. exp. Psychol.*, 1960, 60, 172-179.

8. SIEGEL, P. S., & FOSHEE, J. G. The law of primary reinforcement in children. *J. exp. Psychol.*, 1953, 45, 12-14.

9. SPIKER, C. C. The effects of number of reinforcements on the strength of a generalized instrumental response. *Child Develpm.*, 1956, 27, 37-44.

10. ZIMMERMAN, D. W. Durable secondary reinforcement: method and theory. *Psychol. Rev.*, 1957, 64, 373-383.

11. ZIMMERMAN, D. W. Sustained performance in rats based on secondary reinforcement. *J. comp. physiol. Psychol.*, 1959, 52, 353-358.

EFFECTS OF ASSOCIATIVE STRENGTH, GRADE LEVEL, AND INTERPAIR INTERVAL IN VERBAL PAIRED-ASSOCIATE LEARNING[1]

JOHN C. McCULLERS*[2]

University of Texas

Recent work by Castaneda, Fahel, and Odom (1) has led to a list of words for investigating verbal learning phenomena among elementary school age children by means of the paired-associate method. Although these investigators report associative characteristics for a population of fourth-, fifth-, and sixth-grade children, they do not indicate what, if any, differences in association values arise as a function of grade level. The associative strengths between the word pairs reported by the authors might reasonably be expected to be a function of the number of times the child, during his past experience, has formed an association between any two given words. Accordingly, different associative strengths as a function of grade level might be predicted on the basis of the more general factor of years of linguistic experience as well as the more specific factor of formal (classroom) training in reading and the basic language skills. Thus, if associative strength between the word pairs varies as a function of grade level, differences in paired-associate learning would also be expected.

Furthermore, it may be noted that the original study employed only a single temporal interval (2.0 sec.) between the presentation of word pairs. The present study varies the interpair interval at four levels as well as the factors of grade level and association value and should provide infor-

* Institute of Child Development, University of Minnesota, Minneapolis 14.

[1] This study was supported in part by a Public Health Service Predoctoral Research Fellowship (MR-13000) from the National Institute of Mental Health to the author and in part by grants from the Hogg Foundation, the University of Texas, and the National Institute of Mental Health (Grant M-4240) to Dr. Alfred Castaneda, under whose supervision the present research was carried out.

[2] The author wishes to express his appreciation to the directors of the Austin Public Schools, to Mr. Fred Matthys, Principal, Highland Park Elementary School, Austin, Texas, and to Mrs. Leila Fahel of the Office of Child Research, University of Texas, for their assistance and cooperation.

mation regarding either the main or the interactive effects among these three variables. The present study, then, may be viewed as an extension of the Castaneda, Fahel, and Odom investigation with a twofold purpose of providing information relevant to the questions raised above and also of serving as a further check on the usefulness of the word association list described in the initial study.

Method

Subjects

The subjects consisted of 80 elementary school children selected on the basis of grade level; two groups of 40 subjects each were selected at the fourth and sixth grade levels.

Apparatus and Materials

The apparatus used for exposing the words was the Hunter Card Master (Model 340). A 3-sec. anticipation interval during which the stimulus word was exposed preceded a 3-sec. exposure of the response word. There was, depending on the particular experimental condition, either a 2-, 3-, 4-, or 6-sec. interval between the exposures of the pairs of words. The words used to form the experimental lists for both the high and low association conditions were chosen so as to minimize response competition within the lists. This was accomplished by placing the following restrictions on the selection of words from the normative list: no word was selected for use which (a) carried associative strength, as determined from responses in the normative study, for any word in the list other than its associate or which (b) carried associative strength mediated through some neutral word for any other word in the list except its associate. The final lists were, as nearly as possible, made alike in all respects except in the associative strengths of the word pairs. The lists of the word pairs and their association values are given in Table 1.

Table 1

WORD PAIRS AND THEIR ASSOCIATION VALUES (in parentheses)

	Response	
Stimulus	*Group HI*	*Group LO*
LITTLE	SMALL (60)	THIN (2)
ANGRY	MAD (61)	CRUEL (2)
SLOW	FAST (50)	RUN (5)
AFRAID	SCARED (61)	KNIFE ((2)
WARM	HOT (57)	SUMMER (2)
BEAUTIFUL	PRETTY (62)	HOUSE (2)

Procedure and Design

Each *S* was brought individually into the experimental room and seated in front of the apparatus. The apparatus was located on a table at a height which facilitated *S*'s viewing of the exposed words.

As a precautionary measure to ensure that *S*'s rate of learning would not be hindered by difficulties in reading and understanding, *E* began by asking *S* to read aloud a list of 36 words which were prepared in alphabetical order and which included the 12 words that were to be used in the task. The 24 words on the list which were not used in the task were selected from the normative group of words in such a way that none of these words had association values indicated for any of the 12 words used in the task. The *E* corrected pronunciation and asked *S* to repeat the list in case of one or more reading errors.

Since the instructions vary slightly from those of the original study, they are presented verbatim as follows:

> Now I will read the instructions to you. This is a game to see how well you can remember. This is how it goes. Here on this side (*E* points to the left aperture) a little window will open up, and you will see a word. A few seconds after this window opens, the window on this side (*E* points to right aperture) will open and you will see a *second* word. Your job is to guess what the *second* word is before its window opens. You say the word out loud so that I can hear you. Some of the words on the list that you read will be used in the machine. I had you read the list so that I would know that you could read anything that turns up in the machine. Do you think you understand how it goes now? Fine; now we are ready to start. Remember, you are to say out loud the second word which is going to appear in the second window *before* the window opens. In the beginning you will make a lot of mistakes, but as you learn how it goes, you won't make so many mistakes.

The subjects at each grade level were randomly assigned either to a group which was presented six pairs of words all having high association value (Group HI) or to a group which was presented six word pairs having low association value (Group LO).

Four conditions differing in the length of the interpair interval were used. The interpair interval was considered to be that interval which occurred between the exposures of separate pairs of words and in which both windows of the apparatus were closed. Interpair intervals of 2, 3, 4, and 6 sec. were employed. A $2 \times 2 \times 4$ factorial design consisting of 16 cells or conditions was thus generated by the variables of grade level, association value of the words, and the length of the interpair interval. Five *S*s were assigned to each of the eight conditions within each grade level. The groups were fairly evenly balanced with respect to sex.

After giving the instructions to *S*, *E* seated himself to the side of *S* in such a way as to be able to see the windows of the apparatus and did not further interact with the *S*. Each child was carried to a criterion of one

errorless trial or to a maximum of 15 trials. A trial consisted of one exposure of each of the six word pairs. In order to prevent the word pairs from being learned serially, three different orders of exposing the pairs were employed.

Results and Discussion

The mean number of trials to criterion was computed for the *Ss* in each condition defined by grade level, association value, and interpair interval. This information, which is presented graphically in Figure 1, reveals a tendency toward a grade-level effect. As may be seen, with the exception of the low association condition at the 6-sec. interval, sixth graders reached criterion earlier than fourth graders at every level of interpair interval for both conditions of association. Inasmuch as all *Ss* reached the criterion of one errorless trial, a $2 \times 2 \times 4$ factorial analysis of variance was performed using the number of trials to criterion as each individual's score.[3] The results of this analysis indicated that the factor of grade level did not achieve statistical significance ($F = 3.15$, $df = 1, 64$). The *Ss* for this study were selected at the fourth and sixth grade levels, which constituted the extremes in grade level of the population studied by Castaneda, Fahel, and Odom, on the assumption that any grade-level effect which might be present would be uncovered. Odom (2) also failed to find a significant effect of grade level in a study employing word pairs from the same normative list. The results of these two studies, when considered together, tend to support the notion that for paired-associate tasks among fourth, fifth, and sixth graders, the word association values reported by Castaneda, Fahel, and Odom are not significantly influenced by grade level.

Figure 1 also shows the sharp differences obtained in rate of learning as a function of the association value of the words used. This difference between the HI and LO groups proved to be significant well beyond the .001 level ($F = 26.31$, $df = 1, 64$). This finding furnishes additional experimental verification of the general utility of the original word association list as a means of varying the parameter of associative strength.

Interpair interval proved to be a nonsignificant factor with the particular intervals chosen. It should be pointed out, however, that while the interpair interval normally employed in adult studies is of a duration of less than 1 sec., the shortest interval used in the present study, due to physical limitations of the apparatus, was 2 sec. Thus, all of the interpair intervals employed in the present study were equal to or longer than the one employed by Castaneda, Fahel, and Odom. Therefore, while it may be concluded that increasing the interpair interval by as much as 4 sec. does not appear to affect significantly rate of learning, the possibility remains that shortening the length of this interval below 2 sec. may produce a significant effect. In line with this, interpair interval might also be expected to interact with

[3] The same analysis performed using errors as scores yielded essentially the same results.

anticipation interval and perhaps the duration of time in which the response item is exposed.

Figure 1 further reveals a slight tendency toward a triple interaction effect. There is also a tendency, particularly in the fourth graders, for poorer performance to be associated with increasing length of the interpair interval. The statistical analysis, however, indicated that none of the interactive effects approached significance.

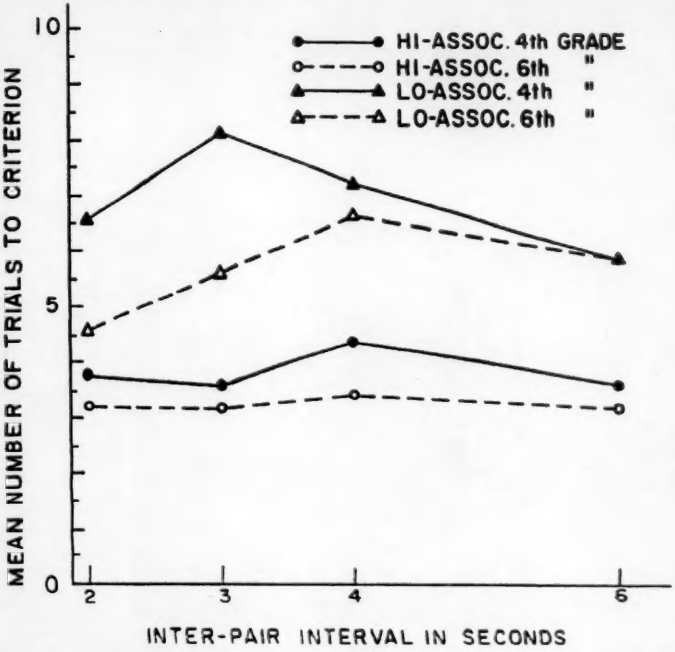

FIGURE 1—Mean number of trials to criterion plotted separately for each condition.

SUMMARY

The present study investigated verbal paired-associate learning among 80 school children at the fourth and sixth grade levels by means of the anticipation method. The variables examined were grade level, association value of the word pairs, and interpair interval. Two levels of association value were employed, a high and a low association condition, and the interpair interval was studied at four levels. The only factor reaching statistical significance proved to be the main effects factor of association value. Some

discussion and evidence are offered concerning the negative findings with regard to grade level.

References

1. Castaneda, A., Fahel, L. S., & Odom, R. Associative characteristics of sixty-three adjectives and their relation to verbal paired-associate learning in children. *Child Develpm.*, 1961, 32, 297-304.
2. Odom, R. Relation of association-value, competition and grade level to verbal paired-associate learning in children. Unpublished master's thesis, Univer. of Texas, 1960.

BEHAVIOR OF YOUNG CHILDREN IN A TWO-CHOICE GUESSING PROBLEM[1,2]

WILLIAM KESSEN* *and* MARION LORD KESSEN
Yale University

Humphreys' basic research (6) on conditions of uncertain outcome has been followed up along several interrelated lines—partial reinforcement studies (*see* 7), tests of mathematical models of probability learning (*see* 3), and studies of cognitive structure (e.g., 4, 5). Several recent studies (8, 10, 11) have extended the range of these settings to include the behavior of children. Convincing regularities have been demonstrated in the asymptotic behavior of animals and men in situations of uncertain outcome, and recent research has turned to an investigation of the variables other than degree of uncertainty which control asymptotic performance (e.g., 9) and to the study of transitional probabilities in response patterns (e.g., 12).

The present study developed from an interest in the following questions: (a) Do young children show the regularities in performance which have been shown in adults? (b) Does the behavior of children in an uncertain-outcome situation vary with prior experience in such situations? (c) Does prior experience have different effects at different ages? (d) What are the patterns of two-choice guessing in young children; i.e., what do the "rules of solution" or "strategies" look like when expressed in terms of transitional probabilities of responding? The research was designed to point up developmental variation and its implications for theories of behavior in situations of uncertain outcome; no formal mathematical model for probability learning, dealing either with asymptotic performance or with sequential dependencies, was under test.

METHOD

Subjects

Seventy-two children from two age-graded classes in each of two nursery schools served as *Ss*.[3] Within each age group (called group A and group B below), children were assigned at random to one of four experimental

* Department of Psychology and Child Study Center, Yale University, New Haven, Connecticut.

[1] This research was supported in part by USPHS Research Grant M-1787 and was completed while the first author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, Stanford, California.

[2] The authors want to thank Richard C. Atkinson, John Gilbert, and George Mandler, who helped in the preparation of this report in several unduplicable ways.

[3] Miss Omwake and her staff at the Yale University Child Study Center Nursery School and Mrs. Norman and her staff at the Yale University Divinity School Nursery School, both of New Haven, Connecticut, gave freely of their time and convenience in providing *Ss* for the research.

groups (called 50/50, 50/70, 70/50, and 70/70 below). Because no significant variation was found between the children from different schools, their data are reported together. Table 1 shows the distribution of *Ss* by age group, school, and sex. The median age of *Ss* in group A was 3 years, 7 months; the median age of *Ss* in group B was 4 years, 5 months. Furthermore, accepting the nursery schools' division into classes left only three *Ss* of group A over 4 years old and only five *Ss* of group B under 4 years old.

TABLE 1

DISTRIBUTION OF SUBJECTS BY AGE GROUP, SEX, AND SCHOOL

	Group A Median Age 3-7		Group B Median Age 4-5	
	Boys	*Girls*	*Boys*	*Girls*
School 1	9	7	8	11
School 2	9	11	7	10

Materials

Blank-faced playing cards were printed with one of two designs—a solid black rectangle 1 inch by 2 inches or a solid red rectangle of the same dimensions. The cards were arranged in three packs of 50 cards each; two packs had an over-all distribution of 25 cards with black rectangles and 25 cards with red rectangles, and the third pack had 35 cards with black rectangles and 15 cards with red rectangles. A further restriction was imposed: for each block of 10 cards, the over-all distributions were the same (i.e., either five black and five red cards or seven black and three red cards), a restriction which limits the length of runs of one design compared to packs drawn without such restrictions. The sequence of cards with black (B) and red (R) rectangles in the 35-15 pack was BRBRBBRBBB BBBRBRBBBR BRBBBRBRB RBRBBBBRBB BBRRBBRBBB. The sequences of the two packs of 25 B and 25 R were RRBBRBBBRR RBRBRBBBRR RBRBBBRBRR *RBRRRBBBRB* BBRRRBBBRR and BRBBBRBRRR BRBRRBRBBR RBBRRBBBRR BBRRBRBBRR BRRBBRRBBR.

Procedure

Each *S* was taken by *E* from his nursery school classroom to another room in the same building and presented with the task individually. With *E* seated on one side of a table and *S* seated on the other, *S* was told: "We are going to play a game. I have some cards here. (*E* exposed two cards illustrating the red and black rectangle). The cards are like these. All of the cards are either red or black. Do you know what color this is? (*E* indicated one of the cards). And what color is this? (*E* indicates the other card). Now what I want you to do is tell me what color you think the next card

in the pack will be. What color do you think the next card is going to be?" These instructions were varied slightly to suit the child, but invariably contained the notions of game, of predicting the color of the next card, and an inquiry to determine the child's color recognition. *S* then said either "Red" or "Black," and the first card was turned so that he could see it. He then said either "Red" or "Black," the next card was exposed, and so on through 50 cards. After each card had been shown to *S*, *E* put it on top of the preceding card. After the first 50 cards had been shown, *E* took a pack of cards and said to *S*: "Now we have some more cards. What color do you think this one will be?" The second pack of 50 cards was then run through in the same way as was the first.

Most children responded well to the task and cooperated with *E*, though with some reluctance on occasion. For the children in this study and with the materials used, 100 guesses seemed close to *S*s' limits of interest and attention.

The four experimental groups varied in the composition of the packs of cards to which they were exposed. Group 50/50 predicted and saw exposed a pack made up of 25 cards with black rectangles and 25 cards with red rectangles on both the first and second runs of 50 cards. Group 50/70 predicted and saw exposed a pack of 25 B and 25 R on the first run and a pack of 35 B and 15 R on the second run. Group 70/50 was exposed to a pack of 35 B and 15 R on the first run and a pack of 25 B and 25 R on the second run. Group 70/70 was exposed to a pack of 35 B and 15 R on both the first and second runs. Thus, for each age group, one experimental group had 100 exposures of a half black - half red composition, one had 100 exposures of a 70 per cent black - 30 per cent red composition; one had 50 exposures of half black - half red composition and was shifted to 50 exposures of 70 per cent B and 30 per cent R; and finally, one experimental group at each age level first had 50 exposures of 70 per cent B and 30 per cent R followed by 50 exposures of half black - half red composition.

Results

Level of Performance

The performance of groups A and B over 10 blocks of 10 responses each is shown in Figure 1 and Figure 2. The performance of groups 50/50

Table 2

ANALYSIS OF VARIANCE OF BLACK RESPONSE, TRIALS 41–50

Source	*df*	*MS*	*F*	*p*
Pack composition (P-c)	1	25.68	4.93	$< .05$
Age (A)	1	2.35	< 1	
P-c × A	1	.35	< 1	
Within groups	68	5.20		

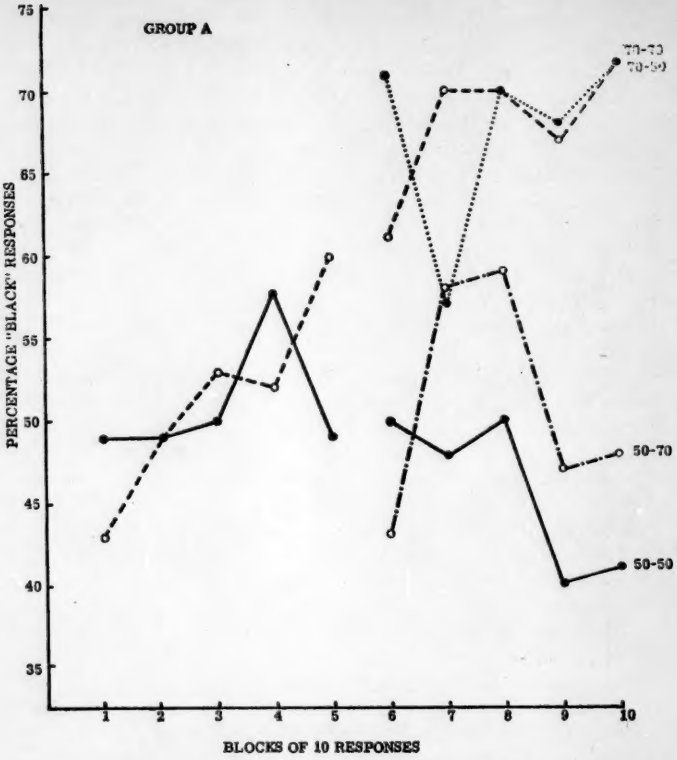

FIGURE 1—Percentage of "Black" responses for younger children (group A). Dashed curve for first five blocks shows responses to 35-15 pack; solid line shows responses to 25-25 pack.

and 50/70 is shown in a single line for the first 50 responses, as is the performance of groups 70/50 and 70/70. At the end of the first fifty guesses and card exposures, both age groups appear to be responding differentially to the 50 per cent and 70 per cent sequences of cards, with the absolute difference in percentage of "Black" response being larger for the older than for the younger children. Table 2 presents an analysis of variance of performance on trials 41 to 50 and indicates that the composition of the packs made a significant difference in performance, but that neither age nor the interaction of age with pack composition approaches statistical significance.

After 100 guesses and card exposures, the groups that had two runs of the same pack (group 50/50 and group 70/70) show different perform-

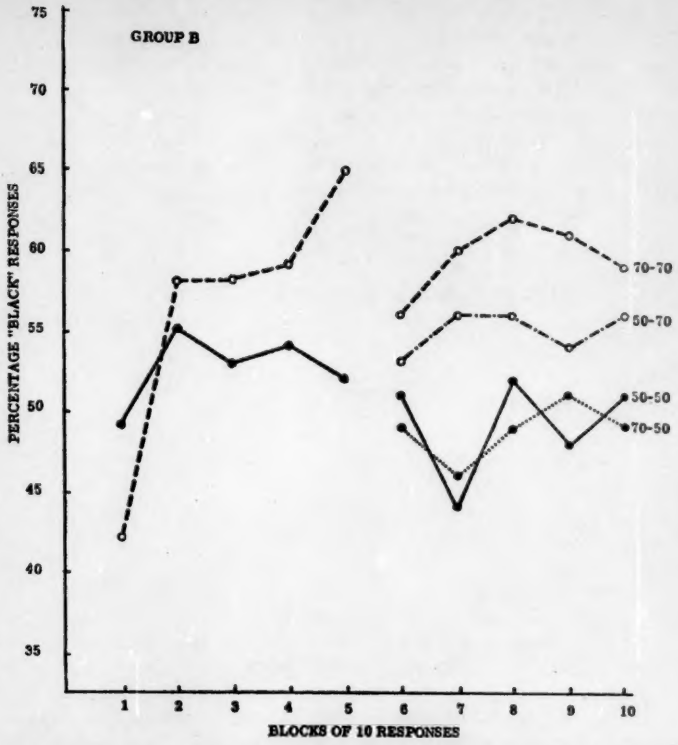

FIGURE 2—Percentage of "Black" responses for older children (group B). Dashed curve for first five blocks shows responses to 35-15 pack; solid line shows responses to 25-25 pack.

ance levels ($p < .01$). It is the behavior of the 50/70 and the 70/50 groups that is of particular interest, however, and particularly among the younger children. As can be seen from Figure 2, the children in group B who were shifted from one pack to another responded to the shift by changing their guessing behavior to conform with the composition of the *second* pack (i.e., they responded "Black" more often to the 35-15 pack than to the 25-25 pack), although the difference between group 70/50 and group 50/70 is less than that between group 70/70 and group 50/50, presumably as a consequence of varying prior experience. The children in group A, as can be seen in Figure 1, behaved in a paradoxical way to the shift in packs. Both the 70/50 and the 50/70 groups performed in a way appropriate to the *first* pack; thus, the children who guessed and saw exposed a pack of cards half

with black rectangles and half with red and then were shifted to a pack in which 70 per cent of the cards were black continue to guess "as if" the cards were still half and half; a similar persistence was shown by these younger children shifted from the 35 B - 15 R pack to the 25 B - 25 R pack. Table 3 describes these results in an analysis of variance of performance during trials 91 to 100. Composition of the pack used for the first 50 responses has a statistically significant effect and, as the figures suggest, the *interaction* between age and the composition of the first pack was statistically significant.

TABLE 3

ANALYSIS OF VARIANCE OF BLACK RESPONSE, TRIALS 91–100

Source	*df*	*MS*	*F*	*p*
Pack composition (P-c) 1st 50	1	36.13	6.71	$< .05$
Pack composition (P-c) 2nd 50 ...	1	5.01	< 1	
Age (A)	1	4.01	< 1	
P-c 1st 50 × P-c 2nd 50	1	.01	< 1	
P-c 1st 50 × A	1	33.35	6.20	$< .05$
P-c 2nd 50 × A	1	.69	< 1	
P-c 1st 50 × P-c 2nd 50 × A	1	3.06	< 1	
Within groups	64	5.38		

In summary, *Ss* responded differentially to two sets of cards, one containing half black and half red markings and the other containing 70 per cent black and 30 per cent red markings, after 50 cards had been exposed. When a second run of 50 exposures was made with the same packs, this result was repeated. In the case of children shifted from one pack to another (i.e., from 50 per cent to 70 per cent and vice versa), the younger children showed a strong tendency to respond in a way appropriate to the distribution of the first-seen pack. Stated in another way, the younger children are less willing to give up an operating rule or response tendency once it has been established.

Transitional Probabilities

Developmental variation of a somewhat different order can be seen in an examination of transitional probabilities of responding "Black" or "Red." How do the patterns or "rules" of response sequence vary with age or prior experience? Suppes and Atkinson (12) have presented an adaptation of a technique of analysis introduced by Anderson and Goodman (2) which is appropriate to this question. Each response of each child (i.e., on every trial n), whether "B" or "R," was tabulated according to whether on the preceding trial (trial $n-1$) his response had been "B" and correct (B+), "R" and correct (R+), "B" and incorrect (B−), or "R" and incorrect (R−). From this basis, the likelihood of saying "B" or "R" after each of these four prior events was computed. Table 4 shows the resulting tabula-

tion of transitional probabilities for the last 40 responses of each run for both age groups and both pack compositions. Thus, Table 4 presents a summary statement of *Ss*' tendency to respond "B" or "R" (a) as this tendency relates to how they responded on the immediately preceding trial and (b) as it relates to whether or not *Ss* had been correct on their immediately preceding guess. For example, the figures in the upper lefthand corner of Table 4 show that, for the younger group of children seeing the 25-25 pack for the first time, the response "B" occurred 56 per cent of the time after the response "B" had occurred on the preceding trial and was correct. The same tabulation shows that the response "R" occurred only 39 per cent of the time after "R" had occurred on the preceding trial and been wrong. The behavior of the younger and older groups can be compared by examining the probability values shown in the top half of the table (group A) with those shown in the bottom half (group B).

TABLE 4

PATTERNS OF RESPONSE SEQUENCE (TRANSITIONAL PROBABILITIES) FOR ALL GROUPS IN ALL CONDITIONS
(Trials 11 to 50 and 61 to 100 are represented.)

	First Run				Second Run							
					25-25 Pack				35-15 Pack			
On Trial $n-1$	25-25 Pack On Trial n		35-15 Pack On Trial n		*After* 25-25 On Trial n		*After* 35-15 On Trial n		*After* 25-25 On Trial n		*After* 35-15 On Trial n	
	B	R	B	R	B	R	B	R	B	R	B	R
Group A												
B+ . .	.56	.44	.60	.40	.43	.57	.82	.18	.54	.46	.73	.27
R+ . .	.53	.47	.65	.35	.51	.49	.39	.61	.69	.31	.97	.03
B− . .	.59	.41	.68	.32	.53	.47	.84	.16	.79	.21	.65	.35
R− . .	.39	.61	.32	.68	.33	.67	.29	.71	.31	.69	.54	.46
Group B												
B+ . .	.35	.65	.51	.49	.34	.66	.29	.71	.46	.54	.49	.51
R+ . .	.74	.26	.86	.14	.68	.32	.60	.40	.80	.20	.91	.09
B− . .	.40	.60	.45	.55	.24	.76	.26	.74	.34	.66	.62	.38
R− . .	.73	.27	.71	.29	.70	.30	.81	.19	.66	.34	.64	.36

These comparisons were subjected to the statistical analysis suggested by Suppes and Atkinson, with the outcomes shown in Table 5. In all six over-all comparisons of the transitional behavior of the two age groups (e.g., first run of 25-25 pack for group A against first run of 25-25 pack for group B), significant differences appeared. Table 5 also shows the significance of the differences in behavior of the two age groups for *each preceding event* (i.e., B+, B−, R+, R−). Thus, the difference between the behavior of group A and that of group B after responding "R" on the preceding trial and being wrong (row R−) is highly significant in five of the six situations.

For the table as a whole 15 of the 24 comparisons between age groups are significant beyond the .01 level of chance expectation, three are beyond the .05 level, and six fall short of statistical significance.

TABLE 5

CHI SQUARE VALUES ($df = 1$) FOR DIFFERENCES IN TRANSITIONAL PROBABILITIES OF RESPONSE BETWEEN GROUP A AND GROUP B

	First Run		Second Run			
	25-25	35-15	25-25 *after* 25-25	25-25 *after* 35-15	35-15 *after* 25-25	35-15 *after* 35-15
B+	16.82**	4.75*	1.58	59.61**	1.62	20.88**
R+	15.73**	11.12**	5.52*	6.54*	1.63	1.08
B−	14.55**	11.65**	14.71**	71.59**	22.44**	.21
R−	40.29**	63.04**	26.49**	40.55**	28.71**	1.62

* $p < .05$.
** $p < .01$.

Examination of Table 4 suggests several summary conclusions. First, the older children were more stable in the application of their operating principles to packs of different composition. For group B, the three sets of transitional probabilities for the 25-25 pack show marked similarity, as do the three sets of transitional probabilities for the 35-15 pack. This regularity is not readily apparent in group A. Second, the older children alternated their responses more frequently than did the younger children. This difference can be seen throughout Table 4 by comparing the likelihood that "B" will follow "R" and that "R" will follow "B" with the likelihood that two "B's" or two "R's" will occur consecutively. In 23 of the 24 possible comparisons shown in Table 4, the younger children follow "B" with "B" and "R" with "R" at higher probability than do the older children and follow "B" with "R" and "R" with "B" at lower probability. Furthermore, if it is recalled that 18 of these comparisons are statistically significant (the single reversal is not), the age difference in the tendency to alternate would seem to be established. Third, the relative tendency among the older children to alternate is particularly in evidence when they have made an incorrect prediction. Compare, for example, the response of group A to the first run of the 25-25 pack when incorrect on the preceding trial with the corresponding pattern of group B. After saying "B" and being wrong, the younger children shifted their response to "R" 41 per cent and repeated "B" 59 per cent while the older children shifted their response to "R" 60 per cent and repeated "B" 40 per cent; correspondingly, after saying "R" and being wrong, younger children shifted 39 per cent and repeated 61 per cent while older children shifted 73 per cent and repeated 27 per cent. This tendency for the younger children to repeat when wrong and for the older children to shift

response when wrong runs through the data. To assess this differential tendency, all *Ss* were scored over trials 11 to 50 as to whether they alternated more frequently when correct on the immediately preceding trial or when incorrect on the immediately preceding trial. Twenty-three of the 36 younger children alternated more often when correct; only 14 of the 36 older children did so. This difference is statistically significant at the .05 level of confidence. As was the case in the analysis of level of performance, the younger children show reluctance in giving up the repetition of a guess even in the face of disconfirming evidence.

In summary, the differences between the two age groups in their "rules of guessing" are large, statistically significant, and regular in regard to the tendency to alternate or to repeat when wrong on the preceding guess.[4] From the data in hand, it cannot be determined whether these differences represent a systematic "development" toward adult patterns of behavior in the two-choice setting. Only the analysis of transitional response probabilities over a wider age range will permit statements about general trends in the effects of age and its correlates on guessing behavior.

These results furnish a demonstration of differences in behavior in the two-choice situation which cannot be dealt with solely by a consideration of the stimulus materials to which the *Ss* were exposed. The differences between age groups point up the interactive determination of behavior in situations of uncertain outcome. Not only is variation in behavior related to the materials used, but variation is also related to the patterns of guessing that *Ss* bring to the task.[5] Under the conditions used in the present study, there appears to be a dramatic shift in the strategy of predicting uncertain outcomes near the end of the fourth year of life.

Summary

Seventy-two preschool children in two age groups predicted the sequential occurrence of cards marked with either a black (B) or a red (R) rectangle. The behavior of four experimental groups, in a two-by-two design for each age group, was studied; one group predicted a pack of 50 cards that was 50 per cent B - 50 per cent R and then predicted a pack of 50 cards 70 per cent B - 30 per cent R, one group predicted the packs in the reverse order, and two groups predicted either 50 per cent B - 50 per cent R or 70 per cent B - 30 per cent R packs in both runs of 50 cards each.

The children responded differentially to the 50 per cent and 70 per cent packs. The older children who were shifted from the 50 per cent pack to the

[4] The examination of the influence of events more remote than the preceding trial can be properly conducted only with longer series of responses than could be obtained in the present study.

[5] Anderson (1), in a study of the effect of first-order conditional probabilities, has presented a related argument for the closer study of individual differences in patterns of alternation and repetition.

70 per cent pack or from the 70 per cent pack to the 50 per cent pack showed a tendency to shift their guessing behavior appropriately; the younger children who were shifted from one pack to another showed a paradoxical tendency to continue guessing in a way appropriate to the *first* run of cards.

An analysis of transitional probabilities of responding showed clear differences between the two age groups. The younger children were more likely to repeat the response made on the preceding trial than were the older children, and this difference was more clearly seen when the guess made on the preceding trial was incorrect.

The results were interpreted as demonstrating the interaction of stimulus materials and strategies of prediction in the determination of behavior in situations of uncertain outcome.

References

1. Anderson, N. H. Effect of first-order conditional probability in a two-choice learning situation. *J. exp. Psychol.*, 1960, 59, 73-93.
2. Anderson, T. W., & Goodman, L. A. Statistical inference about Markov chains. *Ann. math. Statist.*, 1957, 28, 89-110.
3. Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: a study of a science*. Vol. 2. McGraw-Hill, 1959. Pp. 380-491.
4. Feldman, J. An analysis of predictive behavior in a two-choice situation. Unpublished doctoral dissertation, Carnegie Institute of Technology, 1959.
5. Goodnow, J. J. Determinants of choice-distribution in two-choice situations. *Amer. J. Psychol.*, 1955, 68, 106-116.
6. Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, 25, 294-301.
7. Jenkins, W. O., & Stanley, J. C., Jr. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, 47, 193-234.
8. Messick, S. J., & Solley, C. M. Probability learning in children: some exploratory studies. *J. genet. Psychol.*, 1957, 90, 23-32.
9. Siegel, S., & Goldstein, D. A. Decision-making behavior in a two-choice uncertain outcome situation. *J. exp. Psychol.*, 1959, 57, 37-42.
10. Stevenson, H. W., & Weir, M. W. Variables affecting children's performance in a probability learning task. *J. exp. Psychol.*, 1959, 57, 403-412.
11. Stevenson, H. W., & Zigler, E. F. Probability learning in children. *J. exp. Psychol.* 1958, 56, 185-192.
12. Suppes, P., & Atkinson, R. D. *Markov learning models for multiperson situations*. Stanford Univer. Press, 1960.

EFFECTS OF DELAYED REINFORCEMENT ON ASSOCIATIVE AND INCENTIVE FACTORS[1]

ROBERT WARE *and* GLENN TERRELL*
University of Colorado

Numerous experiments both at the human and infrahuman levels have demonstrated that acquisition is in part a function of the temporal nearness of reinforcement to the response (8, 9, 10, 11). Some effort has been made to determine why immediate reinforcement results in more efficient learning. Perin (9) offered a generalized extinction-inhibition hypothesis to account for the detrimental effect of delayed reinforcement on acquisition. Logan (8) and Lipsitt and Castaneda (7) theorized that the time interval separating reinforcement from reaction affects J, which is positively related to the classical conditioned fractional anticipatory goal response, r_g—s_g. Logan demonstrated that delayed reinforcement affects H in rats. Terrell and Ware (11) suggested that both associative and incentive factors are affected by the time interval separating reinforcement from reaction. In the Terrell and Ware experiment, a discrimination learning study, it was reported on the basis of casual observations that *S*s (a) appeared to take less time to decide which one of several stimulus objects to choose and (b) were apparently more interested in the tasks and pushed the response buttons more quickly when performing under an immediate than a delayed reinforcement condition.

It is believed that, in a discrimination learning study, the time interval between the presentation of a stimulus complex and the *onset* of *S*s' response can reasonably be conceived of as an index of the strength of the

* Department of Psychology, University of Colorado, Boulder.

[1] This investigation was supported by a research grant, M-2063, from the National Institute of Mental Health of the National Institutes of Health, United States Public Health Service. Grateful acknowledgment is due Nevin J. Platt, Melvin Wiesley, Mrs. Earl W. Almgren, and all the teachers of the Burke Elementary School, Fairview District, Boulder, Colorado, without whose assistance this experiment could not have been performed. The writers wish to express appreciation to Miss Colette C. Sanville, graduate student, for her assistance in collecting the data of this study.

Child Develpm., 1961, 32, 789-793.

association between stimulus and response components in the learning situation. Hulicka (5) employed a similar definition of the associative variable. Also, in an experiment which requires *Ss* to depress a response button, it is believed that the *speed* with which they *push* the button can be conceived of as an index of *Ss'* incentive to perform the task. Employing these two latencies, the present study investigated the relative effects of immediate and delayed reinforcement on associative and incentive factors in two-stimulus size and form discrimination learning involving children.

Method

Subjects

The *Ss* were 20 children: 10 kindergartners, four girls and six boys, and 10 first graders, five girls and five boys. *Ss* were chosen at random from among 46 first graders and 26 kindergartners. The school from which *Ss* were drawn consists of children predominantly from middle class homes.

Materials

The basic materials used in this experiment have been described in detail elsewhere (11). The stimuli were two pairs of three-dimensional geometric objects: for task 1, the size problem, two cube-boxes 2 and 4 sq. in., respectively; for task 2, a form problem, a sphere and a pyramid. The apparatus consisted of a background board with two push-button mounts 6 mm. in height and a panel board locked onto the rear of and at right angles to the background board. On each trial the stimulus set was placed onto the background board immediately in back of the push-button mounts. A reinforcement light was mounted on the panel board. The light went on immediately following *Ss'* responses for the immediate condition and after a seven-second delay for the delayed reinforcement treatment. In order that measurements of association and incentive could be obtained, Standard electric and Hunter electronic timers were added to the basic apparatus. Mounted under the front edge of the background board was a wooden bar containing a slot which housed a microswitch and a 13.5 by 13 in. plywood shield. The Standard and Hunter timers measured reaction time to 1/100 and 1/1000 of a second, respectively. The circuits were so arranged that removal of the shield activated the microswitch which in turn activated the Standard timer. A 1 mm. depression of either of the push buttons broke the circuit to the Standard timer and simultaneously activated the Hunter timer. The Hunter timer was deactivated when the stem of the push button was depressed its full distance by *S*. Measurements of associative and incentive variables were obtained on the Standard and Hunter timers, respectively.

Design and Procedure

Ss were randomly assigned to two groups of 10 children, each containing five kindergartners and five first graders. Group I *Ss* were given the size task

under the delayed reinforcement condition and the form task under the immediate reward condition. For group II *Ss*, the task-reinforcement combinations were the reverse of those for group I. *Ss*, tested individually, received the following instructions: "This is a game where we play with these things (*E* points to the four stimulus objects). You are supposed to choose one of the two things up here (*E* points to the background board) and push the button in front of the one you choose. If you choose the correct one, the light will go on. If you choose the wrong one, the light will not go on." The experimenter informed *S* that the shield would be inserted in place before each trial. This made it impossible for *S* to observe *E* exchange stimulus sets between trials. Also, removal of the shield activated the Standard timer. The delay involved in replacing the shield in the slot after each trial was consistent with the reward condition of the just completed trial: Training was continued until *Ss* reached criterion (six consecutive correct responses) on each of the two tasks. For task 1 the correct response was a button-pushing response to the larger of two cube boxes, and for task 2, a response to a sphere instead of a pyramid.

Results

As was predicted, *Ss* learned significantly more quickly under the immediate than under the delayed reinforcement condition ($p = .005$). The Wilcoxon matched-pairs signed-ranks test was used to analyze both latency measures (*see* Table 1). The effects of the reinforcement conditions on both measures were determined at three points in the learning process: (a) the first six trials, (b) at the point where *Ss* reached criterion on the first of the two tasks (the six adjusted trials),[2] and (c) the six criterion trials. For both of the latency measures a raw score for each *S* for each of the two reinforcement categories was determined by summing the latencies of six trials at the three points in the learning process. The associative latencies, with the exception of the first six trials analysis, were significantly shorter under the delayed than under the immediate condition ($p = .44$, .003, and .02 for the first six trials, adjusted trials, and criterion trials, respectively). The reinforcement conditions did not differentially affect the incentive latencies ($p = .50$, .35, and .62 for the first six trials, adjusted trials, and criterion trials, respectively). It is believed that the incentive measure was not sensitive enough to pick up possible differences in the effect of the treatment conditions.

[2] Adjusted scores were determined by the number of trials to the first criterion and the reinforcement condition under which *Ss* first reached criterion. For example, if *S* required 10 trials (exclusive of the six criterial trials) under the immediate and 18 trials (exclusive of the six criterial trials) under the delayed reinforcement condition to learn the correct response, the sum of the six criterial trials (trials 11 to 16, inclusive) under the immediate condition was compared with the sum of trials 11 to 16, inclusive, under the delayed reinforcement condition. If *S* reached criterion first under the delayed reinforcement condition, measurements under the immediate condition were adjusted. This adjustment was made for each *S*.

TABLE 1

WILCOXON MATCHED-PAIRS SIGNED-RANKS ANALYSIS OF ASSOCIATIVE AND INCENTIVE MEASURES

(For each comparison combination, $N = 20$)

Measure	*T*	*p**
Associative		
First Six Trials	84.5	.44
Adjusted Trials	24.0	.003
Criterion Trials	43.0	.02
Incentive		
First Six Trials	105.0	.50
Adjusted Trials	80.0	.35
Criterion Trials	118.0	.62

* Probabilities are for two-tailed tests of significance.

DISCUSSION

Contrary to impressions suggested by casual observations in a previous experiment (11), *Ss* in the present experiment responded more quickly to the task with which delay of reinforcement was associated (even though they learned this less quickly) than to the task with which immediate reinforcement was associated.

The interference of reinforcement delay with the establishment of correct associations may be interpreted in at least two ways: Perhaps delay of reward thwarts problem-solving behavior. The thwarting of goal-directed behavior is frustrating. Numerous studies both at the human and the infrahuman level have shown that frustration (a) impairs performance on relatively complex tasks and (b) is associated with increased muscular tension which produces faster reactions (1, 2, 3, 4, 6). These studies indicate that frustration and its accompanying tension are negatively related to performance. It seems reasonable to suggest that *Ss*, in the present experiment, frustrated because of the delay of an indication of the correctness or incorrectness of their response, (a) reacted in an impulsive, nonreflective way, (b) gave little if any attention to the significant elements of the task at hand, and (c) made use of previous experience with the task in an inefficient way. On the other hand, when reinforcement followed correct responses immediately, *Ss* were apparently more attentive to the meaningful components of the task situation, a fact which led to a quicker strengthening of the correct habits.

Another possible explanation of these results consists of the likelihood that *Ss* became uninterested and consequently bored with a task which